CONCEPTUALIZATION AND MEASUREMENT OF HEALTH FOR ADULTS IN THE HEALTH INSURANCE STUDY: VOL. V, GENERAL HEALTH PERCEPTIONS

JOHN E. WARE, JR., ALLYSON DAVIES-AVERY, CATHY A. DONALD

R-1987/5-HEW

SEPTEMBER 1978

CONCEPTUALIZATION AND MEASUREMENT OF HEALTH FOR ADULTS IN THE HEALTH INSURANCE STUDY: VOL. V, GENERAL HEALTH PERCEPTIONS

JOHN E. WARE, JR., ALLYSON DAVIES-AVERY,
CATHY A. DONALD

Rand

SANTA MONICA, CA. 90406

PREFACE

The Rand Health Insurance Study, funded by a grant from the U.S. Department of Health, Education, and Welfare, is a social experiment being conducted in six sites across the United States to investigate the effects of different health care financing arrangements (differing coinsurance and deductible rates and fee-for-service practice versus prepaid group practice) on the use of personal medical care services, quality of care, satisfaction with care, and health status. Some 8000 people in 2750 families are enrolled in the experiment for periods of three or five years; health status is assessed for each person on entering the experiment, annually during the experiment, and on leaving.

Developing reliable and valid measures for assessing enrollee health status was a prerequisite to examination of the effects of health care financing on health status in the Health Insurance Study. The volumes that constitute Rand report R-1987-HEW (see below) contain detailed information on the conceptualization and measurement of the health status of adults (age 14 and older) in terms of physical, mental, and social health and general health perceptions. They also present data on the health status of adults upon enrollment in the experiment at the first site (Dayton, Ohio) and revisions made in measures of health status for repeated use in Dayton and other study sites. Measurement of physiologic health is discussed in Robert H. Brook, George A. Goldberg, L. Jeff Harris, Kenneth H. Applegate, Marc Rosenthal, and Kathleen N. Lohr, *Conceptualization and Measurement of Physiologic Health in the Health Insurance Study*, The Rand Corporation, R-2262-HEW, forthcoming. Measurement of the health status of children (under age 14) enrolled in the experiment is discussed in Marvin Eisen, Cathy A. Donald, John E. Ware, Jr., and Robert H. Brook, *Conceptualization and Measurement of Health for Children in the Health Insurance Study*, The Rand Corporation, R-2313-HEW, forthcoming.

The eight volumes in the R-1987-HEW series, which has the overall title *Conceptualization and Measurement of Health for Adults in the Health Insurance Study*, include:

John E. Ware, Jr., Robert H. Brook, Kathleen N. Williams, Anita L. Stewart, and Allyson Davies-Avery, *Vol. I, Model of Health and Methodology*, R-1987/1-HEW.

Anita L. Stewart, John E. Ware, Jr., Robert H. Brook, and Allyson Davies-Avery, *Vol. II, Physical Health in Terms of Functioning*, R-1987/2-HEW.

Shawn A. Johnston, John E. Ware, Jr., Allyson Davies-Avery, and Robert H. Brook, *Vol. III, Mental Health*, R-1987/3-HEW.

Cathy A. Donald, John E. Ware, Jr., Robert H. Brook, and Allyson Davies-Avery, *Vol. IV, Social Health*, R-1987/4-HEW.

John E. Ware, Jr., Allyson Davies-Avery, and Cathy A. Donald, *Vol. V, General Health Perceptions*, R-1987/5-HEW.

John E. Ware, Jr., Robert H. Brook, and Allyson Davies-Avery, *Vol. VI, Analysis of Relationships among Health Status Measures*, R-1987/6-HEW.

William H. Rogers, Kathleen N. Williams, and Robert H. Brook, *Vol. VII, Power Analysis of Health Status Measures*, R-1987/7-HEW.

Robert H. Brook, John E. Ware, Jr., Allyson Davies-Avery, Anita L. Stewart, Shawn A. Johnston, Cathy A. Donald, William H. Rogers, and Kathleen N. Williams, *Vol. VIII, Overview*, R-1987/8-HEW.

Volumes I–VII are directed primarily to those who will be using these measures during Health Insurance Study (HIS) analyses and to other investigators who are interested in using or adapting HIS measures for their own research. Volume VIII summarizes the results and conclusions of studies of these HIS measures of health status for a more general audience. Although every attempt was made to write the volumes so that they might be read without reference to others in the series, this was not always possible. The reader is urged to consult the first volume, in particular, as it describes the model of health adopted for use in the Health Insurance Study, the site and sample selection methods, and the methods used to construct health status measures and study their reliability and validity.

Subsequent reports will present results of revised measures of physical, mental, and social health status and general health perceptions currently in use in the Health Insurance Study.

Additional Rand reports and publications discuss other design and measurement issues related to the study. A preliminary report of issues in health status assessment appeared in Arnold I. Kisch and Paul R. Torrens, "Health Status Assessment in the Health Insurance Study," *Inquiry*, Vol. 11, 1974, pp. 40–52.

The experimental design for estimating the effects of financing on demand for care is described in Joseph P. Newhouse, "A Design for a Health Insurance Experiment," *Inquiry*, Vol. 11, 1974, pp. 5–27; and in Joseph P. Newhouse, *The Health Insurance Study: A Summary*, The Rand Corporation, R-965-OEO, March 1974. Features of the design that permit estimation of the effects on utilization behavior attributable solely to participation in the experiment are discussed in Joseph P. Newhouse, Carl N. Morris, Kent H. Marquis, Charles E. Phelps, and William H. Rogers, "Measurement Issues in the Second Generation of Social Experiments: The Health Insurance Study," *Proceedings*, Social Statistical Section, American Statistical Association, 1976.

Carl N. Morris, *A Finite Selection Model for Experimental Design of the Health Insurance Study*, The Rand Corporation, R-1837-HEW (forthcoming) describes the logic and techniques used to determine optimum sample sizes for the Health Insurance Study and to assign individual families to experimental plans.

The first in a projected series of reports dealing with measurement of consumption of medical services in the Health Insurance Study is Kent H. Marquis, *The Methodology Used To Measure Health Care Consumption During the First Year of the Health Insurance Experiment*, The Rand Corporation, R-2126-HEW, August 1977. The application of reliability theory to evaluation of the quality of survey data such as those in the Health Insurance Study is discussed in M. Susan Marquis and Kent H. Marquis, *Survey Measurement Design and Evaluation Using Reliability Theory*, The Rand Corporation, R-2088-HEW, June 1977.

Other methodological issues related to techniques for obtaining precise, unbiased estimates of medical care expenditures are examined in Kent H. Marquis, M. Susan Marquis, and Joseph P. Newhouse, *The Measurement of Expenditures for Outpatient Physician and Dental Services: Methodological Findings from the Health Insurance Study*, The Rand Corporation, R-1883-HEW, April 1976.

SUMMARY

Self-ratings of general health (as well as physical, mental, and social health) are obtained annually in the Health Insurance Study to test hypotheses regarding the effects of differences in coinsurance and deductibles in a comprehensive health insurance benefits package, and of differences in use of medical care services on individual health status. This volume discusses the conceptualization and measurement of general health perceptions, including: (1) a review of the literature on general health rating measures developed before the Health Perceptions Questionnaire, which was selected for use in the Health Insurance Study; (2) results of psychometric studies of the Health Perceptions Questionnaire in general populations other than that enrolled in the Health Insurance Study; and (3) plans for evaluating the measurement of general health perceptions in the Health Insurance Study and for testing experimental hypotheses. Because scales explicitly constructed to measure general health perceptions were not included in the enrollment Medical History Questionnaire in the first Health Insurance Study site (Dayton, Ohio), this volume contains no data on general health perceptions from the Health Insurance Study.

Taken at face value, measures of general health perceptions differ from other health status measures in that they do not specify one or more components of health (physical, mental, or social). Rather, respondents are asked only for an assessment of their "health." In theory, this difference in measurement strategy makes it possible to achieve two important goals. First, general health ratings may constitute one kind of overall health status index if respondents consider all health components when they make their ratings. Second, general health ratings may reflect the objective information people have about their health status as well as their evaluation of that information and may, thereby, help solve the problem of aggregating the two kinds of health status data.

The literature review addressed the issues of reliability and validity of general health perception measures and the likelihood that such measures would contribute information about health status in addition to that available from measures of specific health status components (physical, mental, and social) such as those fielded in the Health Insurance Study. The review included 40 publications pertaining to general health perceptions published during the 25 years before development of the Health Perceptions Questionnaire, as well as those that describe the use of the Health Perceptions Questionnaire in general populations.

The literature review indicated that general health ratings, despite their frequent use, had not been rigorously evaluated in terms of reliability and validity. With few exceptions, general health perceptions were scored from single-item measures, and reliability estimates for these measures were not reported. The limited information about reliability and stability for measures constructed from more than one item suggests that ratings may be reliable enough for group comparisons and stable enough over time to warrant use in a repeated-measures study design.

The validity of general health ratings was rarely addressed explicitly in published research. Reported associations between ratings and other measures of health status and health and illness behavior proved useful, however, in assessing

validity. In most instances, the pattern of associations was observed that would be expected if general health ratings are indeed a measure of *general* health, supporting the notion that people consider more than one component of health status when they rate their health in general. Without reliability data, however, published validity findings could not be evaluated to determine the strength of the associations involved.

Studies of the Health Perceptions Questionnaire (HPQ) indicated that general health perceptions are oriented with respect to time; respondents distinguish among perceptions of what their general health has been, is now, and what they think it will be in the future. On the basis of this finding and scaling studies, the HPQ was constructed to measure the following hypothesized dimensions of general health perceptions: Prior Health, Current Health, Health Outlook, Resistance-Susceptibility to Illness, Health Worry/Concern, and Sickness Orientation.

Scaling studies based on factor analytic and discriminant validity criteria were successful. As required by the scaling method used, retained items in each hypothesized scale were clearly shown to measure the same construct and to measure that construct more than others. The consistency of findings across independent field tests and across populations differing widely in age, level of educational attainment, income, and race supported generalization of conclusions regarding item groupings and the six general health scales they define.

Studies of reliability based on test-retest and internal-consistency methods indicated that the HPQ scales were sufficiently reliable for purposes of group comparisons; single-item scores often were not. Reliability coefficients tended to be lower in disadvantaged populations, although the amount of true score variance defined by scales was sufficient to warrant their use in comparing groups in all populations studied. Comparisons of different kinds of reliability estimates suggested that general health perceptions may be less reliable at any time *and* less stable over time for the disadvantaged.

Factor analytic studies of the construct validity of the HPQ (and additional analyses of data from those studies) supported the validity of scale scores. These results were also useful in determining the meaning that should be assigned to each scale score and support construction of a global health perceptions index using two or more of the six HPQ health scales. Such an index would be very reliable and would capture the majority of reliable variance defined by six HPQ scales. Tradeoffs between the simplicity of a single health perceptions index and loss of predictive validity with one rather than six scores require further study.

Studies of associations between HPQ scales and other health status variables and health and illness behavior also supported the validity of the scales. Although this evidence increased understanding of the meaning of general health ratings, little is known about the clinical significance of scale scores. Correlations among the scales taken into account in evaluating their validities in relation to health and illness behavior suggested that more than one scale may contribute to the prediction of such behavior and that the best predictor may differ across populations.

Before their use in hypothesis testing, the HPQ scales used in the Health Insurance Study will be studied to increase understanding of scores in terms of: reliability; stability over one- to five-year intervals; validity as measures of physical, mental, and social components of health status and predictors of health and illness behavior; and precision. Priorities for analyses of Health Insurance Study data before hypothesis testing were based on results and shortcomings of previous

studies and fall into three categories: (1) research design, (2) comprehensiveness of validity variables, and (3) analytic methods.

Longitudinal data available from the Health Insurance Study will permit prospective studies of the validity of the health perception measures; all previous studies have been cross-sectional. Repeated administrations will also make it possible to determine the validity of scales that differ with respect to time (perceptions regarding prior, current, and future health).

The validity of general health ratings in relation to data obtained from sources other than the respondent will be studied; Health Insurance Study data will permit analyses of these ratings in relation to physician diagnoses, results of comprehensive screening examinations performed by trained technicians, and extent of disability reported in biweekly diaries.

The comprehensiveness of health status data gathered in the Health Insurance Study will make it possible to determine the extent to which general health ratings reflect both objective and subjective health status constructs and specific components of health (physical, mental, and social). Further work will be done during analyses to determine the possibility of constructing a general health perceptions index and its usefulness in hypothesis testing.

ACKNOWLEDGMENTS

CONTENTS

TABLES

I. INTRODUCTION

OVERVIEW

This volume concerns the conceptualization and measurement of health of individuals using self-ratings of general health status (also termed general health perceptions). Examples of general health perceptions measures include the widely used single-item rating of health in terms of "excellent," "good," "fair," or "poor," and endorsement or nonendorsement of such statements as "I have been feeling bad lately."

These measures differ from the health status measures discussed in Vols. II, III, and IV in a number of important ways. First, general health ratings explicitly ask the respondent about personal "health." This is in contrast to operational definitions of health constructs based on descriptions of behavior or feelings that do not use the word "health" (e.g., "ability to walk upstairs" as a measure of physical health; "feeling downhearted" as a measure of mental health). Second, general health ratings do not focus on a single component of health, such as physiologic, physical, mental, or social health. Presumably, they assess a construct underlying specific components of health, and scores on such measures reflect an integrated perception of health status across specific components. Third, general health ratings are subjective rather than objective. They are not described in terms of behavior that can be directly observed, such as a day spent in bed because of poor health. These ratings are designed to assess the subjective as well as objective aspects of health status as perceived by respondents. In this regard, measures of general health perceptions resemble measures of perceived severity and susceptibility to illness in studies of the Health Belief Model (Maiman and Becker, 1974). Finally, if the logic offered by Ware and Karmos (1976) regarding the usefulness of general health ratings is correct, such ratings will be more valid as predictors of use of health care services and of other health and illness behavior than any other kind of measure, including physician assessments. This logic is based on the presumption that people seek care on the basis of a self-assessed need for it.

Because they measure health in a different and potentially more comprehensive way than measures of specific and objective health components, self-ratings of general health have been included in Health Insurance Study (HIS) questionnaires designed to measure individual health status. The first Medical History Questionnaire (MHQ), administered in Dayton, included a four-item scale to measure health worry/concern in Dupuy's (1972) General Well-Being Schedule (GWB) used to measure mental health. Analyses of relationships between this scale and other measures in the HIS battery revealed that this scale was a measure of general health perceptions. The administration and construction of this scale are discussed in Vol. III; its validity in relation to other HIS measures of health status is discussed in Vol. VI.

HIS studies of the GWB general health perceptions measure indicated that although the four-item scale overlaps considerably with measures of physical and mental health (as would be expected if it is a valid general health measure), much of the variance in general health ratings is not predictable from measures of specific

1

health constructs. General health ratings appear to add information to that obtained from objective measures of specific health components. On the basis of such findings and published studies, a separate battery of general health rating items was included in HIS health questionnaires fielded after the Dayton enrollment Medical History Questionnaire and in repeated measures of health status in Dayton. Following an extensive literature review, Form II of the Health Perceptions Questionnaire developed by Ware and Karmos (1976) was selected for use. During research sponsored by the National Center for Health Services Research, this questionnaire was extensively tested in terms of reliability and validity in general populations; results indicated that it would be suitable for self-administration in the HIS and that it is likely to accomplish some of the measurement goals noted above. These results were particularly important because the reliability of measures of general health perceptions had been questioned in the literature, and the validity of the conceptual arguments in favor of their use had not been tested systematically.

CONTENTS AND ORGANIZATION

The next section reviews the literature on general health perceptions. Studies of survey measures of general health perceptions that were developed before the Health Perceptions Questionnaire (HPQ) are discussed first, followed by a more detailed summary of results of studies of the HPQ in non-HIS general populations. The review emphasizes issues that are critical to evaluating the usefulness of general health ratings in comprehensive batteries designed to measure individual health status:

- What has been the content of items used to obtain general health ratings?
- Is there value in distinguishing among perceptions of past, present, and future health?
- What have questionnaires concerned with general health perceptions been used for?
- What data-gathering, scoring, and scaling methods have been used?
- Is a general health rating best viewed as a dichotomy (do persons perceive themselves as either healthy or unhealthy), or is it possible and useful to place persons along a continuum of perceived health?
- What evidence is there that general health ratings are reliable and stable over time?
- Are general health ratings valid in terms of their intended use (as general health measures)? To what extent do they mirror the components of health status (e.g., physical abilities, psychological well-being, and social health)?
- Are scale (multi-item) measures of general health perceptions more reliable and valid than the commonly used single-item ratings of general health?
- What do measures of general health perceptions add to information available from measures designed specifically to assess physical, mental, and social health in terms of describing individual health status?

- What information is available regarding the relationship between demographic and socioeconomic variables and general health perceptions?

Following a summary of the literature review, the planned analyses of HIS data from the HPQ are discussed, and suggestions are offered for future research on measures of general health perceptions.

II. LITERATURE REVIEW: PRE-HPQ
GENERAL HEALTH RATINGS

BACKGROUND

The review of the literature on general health ratings (developed before the Health Perceptions Questionnaire) was based on 40 empirical and theoretical articles published from 1959 to 1977. All articles referred to or described general health perception variables and were selected for more thorough review on the basis of the information they contained.[1] Table 1 identifies the investigators whose work was reviewed and includes information on the purpose of the study, population studied, data-gathering methods, and whether single-item or scale measures were used. (Other information about many of these measures is contained in Reeder, Ramacher, and Gorelnik, 1976.)

Nearly all articles (39 of 40) were empirical—i.e., included previously unreported analyses of general health ratings. Of those, only six based their analyses entirely or primarily on scale scores; 33 of 39 empirical studies used single-item measures.

CONTENT OF QUESTIONNAIRES

The content of items used to measure general health perceptions was similar across questionnaires in the studies reviewed. All items appeared to request a subjective rating of personal health. Respondents were usually asked to choose, from a number of response categories, the one that best described their own health in general. Table 2 illustrates some of the differences in wording of questionnaire items.

Terms used to label general health rating variables also differed among investigators. Table 3 presents some of the labels that have been used; despite variations, they capture the notions of "subjectivity" or "self-assessment" of health. Some investigators used different terms to refer to the same or similar items, suggesting a lack of consensus regarding what these measures yield.

DATA-GATHERING METHODS

Trained interviewers obtained general health ratings in 28 studies; several of these studies combined personal interviews with physician interviews, medical record reviews, or mail questionnaires. Two studies used self-administered questionnaires, one used mail questionnaires, and four did not describe data-gathering methods.

---

[1] Articles were reviewed using a standardized abstraction form that investigated (1) whether original data were reported; (2) the intended use of the measures; (3) the labels used by authors to describe their measures; (4) methods used to gather data; (5) whether verbatim questionnaire items were published or described; (6) the number of independent scores computed; (7) whether variables were defined as single-item or scale scores; (8) types of scale construction methods used, if any; (9) whether reliability data were reported; (10) whether data relevant to validity were reported; (11) variables that were studied in relation to general health ratings; and (12) major findings.

Table 1

SUMMARY OF INFORMATION ABOUT PUBLISHED STUDIES OF GENERAL HEALTH PERCEPTIONS

Investigator(s)	Purpose(s) of Study	N	Population(s) Studied	Data Source/Method of Data Collection	Single-Item or Scale Measure
DiCicco and Apple (1958)	Estimate relationships among health constructs	95	Elderly persons living in low socioeconomic district (60 white; 35 black); ages 65+	Personal interview	Scale
Streib, Suchman and Phillips (1958)	Test reliability and validity of personal health ratings	2875	Elderly persons; ages 65+	Self-administered questionnaire	Scale
Thompson and Streib (1958)	Assess relationship between personal health rating and retirement	1260	Males (477 retired; remainder employed); ages 68-70	Not given	Scale
Schnore and Cowhig (1959-60)	Assess relationship between personal health rating and demographic characteristics	1304[a]	General population urban area probability sample; ages 21+	Personal interview	Single-item
Friedsam and Martin (1963)	Assess relationship between personal and physician ratings of health	87	Sample of white patients of outpatient clinic at large city-county hospital in SW United States; ages 50+	Patient and physician interviews	Scale
Heyman and Jeffers (1963)	Assess relationship between personal assessment and medical evaluation of health over time	182	Community volunteers for geriatrics research program; ages 60-94	Personal interview	Single-item
Maddox (1964)	Correlate personal health ratings with other health and attitudinal variables over time	182	Volunteer subjects in study of human aging; ages 60-94	Personal interview	Single-item
Kessel and Shepard (1965)	Assess relationship between personal health rating and illness behavior	176	Patients of a general practitioner in a middle class suburb of London; ages 11+	Personal interview	Single-item
Desroches, Kaiman, and Ballard (1967)	Assess relationship between personal health rating, age, and self-reported physical and psychological symptoms	22	Elderly subjects in VA domiciliary; mean age 67.09	Personal interview	Single-item
Henley and Davis (1967)	Assess relationship between satisfaction and aspects of one's life situation (e.g., health)	167	Chronically ill ambulatory urban patients; medically indigent; ages 60+	Personal interview	Single-item
Madow (1967)	Compare personal health ratings with data on medical record	5988 (Interviews on stratified sample of 2000)	Sample of members of Kaiser Foundation Health Plan, Southern California; ages 17+	Personal interview and medical records	Single-item

Table 1 (continued)

Investigator(s)	Purpose(s) of Study	N	Population(s) Studied	Data Source/Method of Data Collection	Single-Item or Scale Measure
Shanas, Townsend, Wedderburn, et al. (1968)	Assess relationship of personal health rating and physical incapacity across the elderly populations in various countries	7500 (Approx.)	Cross-national survey of elderly people in Denmark, Britain, U.S. (area probability sampling of persons in private households); ages 65+	Personal interview	Single-item
Hochstim, Athanasopoulos and Larkins (1968)	Compare persons of different race and income groups on a number of health, economic, and social variables	908 (Poverty area) 1672 (Remainder of the city)	Area probability sample of adult residents of Oakland, California; ages 20+ or 16-19 and ever married	Not given	Single-item
Ludwig and Gibson (1969)	Assess relationship between perception of poor health and use of medical services	227	Applicants for social security disability benefits	Personal interview	Single-item
Renne (1970)	Assess relationship between marital satisfaction and social, economic, psychological, and health variables (e.g., evaluation of one's general health)	5163	Probability sample of married adults in a large metropolitan area; ages 20+ or 16-19 and ever married	Mail questionnaires and personal interviews	Single-item
Renne (1971)	Assess relationship between health and marital status/happiness	5373	Probability sample of currently married, separated, or divorced adults in a large metropolitan area; ages 20+ or 16-19 and ever married	Mail questionnaires and personal interviews	Single-item
Pratt (1971)	Assess relationship between health and socioeconomic status	401	Mothers with children aged 9-13 from households having a husband in the residence	Personal interviews	Scale
Berkman (1971)	Assess relationship between mental health and stress factors (e.g., self-evaluated physical health)	6516	Probability sample of household in a large metropolitan area	Mail questionnaires	Scale
Cole and Lejeune (1972)	Evaluate social-psychological determinants (e.g., ability to fulfill normal roles) of personal health rating	3038	Probability and quota samples of welfare mothers and working class black women in New York City and Camden, New Jersey	Personal interview	Single-item

Table 1 (continued)

Investigator(s)	Purpose(s) of Study	N	Population(s) Studied	Data Source/Method of Data Collection	Single-Item or Scale Measure
Fabrega and Roberts (1972)	Evaluate influence of socio-psychological variables (e.g., perceived health status) and use of health services	306	Probability sample of economically disadvantaged Negro urban residents; ages 18+ or married	Personal interview	Single-item
Gaitz and Scott (1972)	Assess relationship between personal health rating and psychological well-being	1441	Stratified sample of noninstitutionalized adult community-dwelling respondents (quota method of sampling); ages 20+	Personal interview	Single-item
Leveson (1972)	Assess relationship between health status and use of services	1219 (Subsample of 323 persons with reported heart conditions)	Sample of residents of housing project; ages 60+	Household survey for non-registrants in a Health Maintenance Survey; data source unclear for registrants	Single-item
Palmore and Luikart (1972)	Assess relationship between health/social factors and life satisfaction	502	Random selection of white persons from the membership lists of a local major health insurance association; ages 46-71	Personal interview	Single-item
Pratt (1972)	Assess relationship between health and marital organization (conjugal power)	783	Analysis based on interviews with random sample of 273 paired husbands and wives and an additional 237 wives; families had to have children aged 9-13	Personal interview	Single-item
Tissue (1972)	Identify factors associated with personal ratings of health in old age	256	Aged welfare recipients; median age: 68 years; 111 men and 145 women	Personal interview	Single-item
Crawford and Hooper (1973)	Examine relationship between menopausal experience and general health patterns, age, sex, identity, and marital relationships	106	43 married women whose first child was about to marry and 63 married women who were expecting the arrival of their first grandchild; median age 49 years; 39% employed outside the home; socioeconomic status skewed toward higher classes	Personal interview	Single-item

Table 1 (continued)

Investigator(s)	Purpose(s) of Study	N	Population(s) Studied	Data Source/Method of Data Collection	Single-Item or Scale Measure
Cutler (1973)	Assess relationship between participation in voluntary associations, life satisfaction, personal rating of health, and socioeconomic status	170	Randomly selected sample of non-institutionalized respondents; ages 65+; median age: 74 years; median family income: $3200; 121 females; 49 males	Personal interview	Single-item
Edwards and Klemmack (1973)	Assess relationship between life satisfaction and 22 variables (including perceived health status)	507	Data from a larger study using a census enumeration district quota sample; ages 45+; predominantly white, Anglo-Saxon, Protestant	Not given	Single-item
Garrity (1973)	Identify factors (e.g. perceived health status) associated with return to work after myocardial infarction	58	White males who had experienced their first myocardial infarctions and had survived at least 6 mos. after hospital discharge; ages 37-74	Interviews, mail questionnaires, hospital records and physician questionnaire	Single-item
Osborn (1973)	Assess relationship between personal health rating and social status	582	Random stratified area probability samples of white, married couples, husbands age 60-64; wives were excluded from the present study	Personal interview	Single-item
Nowlin (1974)	Study relationship between personal health rating and depression over time	157	Ambulatory community volunteers; average age: 72.3 years	Two-day evaluation by Aging Center in Durham, N.C.	Single-item
Renne (1974)	Define and measure social health of individuals	6928	Probability sample of households in an urban area; ages 20+ or 16-19 and ever married	Self-report	Scale
Reynolds, Rushing, and Miles (1974)	Validate Function Status Index	8036	Random sample of persons in two rural economically depressed counties; ages 21+	Personal interview	Single-item
Spreitzer and Snyder (1974)	Study correlates of life satisfaction among older persons	1547	National probability sample of the U.S. population (non-institutionalized); ages 18+	Personal interview	Single-item
Warren (1974)	Assess health of older people	65	Random selection of ambulatory men and women from a low-income housing project ages 60-90; median age: 76.2; 9 males, 56 females	Personal interview	Single-item

Table 1 (continued)

Investigator(s)	Purpose(s) of Study	N	Population(s) Studied	Data Source/Method of Data Collection	Single-Item or Scale Measure
Fine (1975)	Study relationships among health variables; assess characteristics of users and non-users of a transportation program	169	Random sample of urban elderly subjects registered in New York City Half Fare Program for Senior Citizens and random sample of users of a transportation program; ages 65-89 (user group contained a few people under 65)	Personal interview	Single-item
Tornstam (1975)	Study relationship between health and attitudes	469	Random sample of persons in a Swedish town; ages 45-75	Personal interview	Single-item
Bergner, Bobbitt, Pollard, et al. (1976)	Validate Sickness Impact Profile	278	Rehabilitation medicine patients; speech pathology patients; outpatients with chronic problems; group practice enrollees; ages 18-75	Personal interview	Single-item

[a]Weighted sample based on 974 interviews

Table 2

EXAMPLES OF QUESTIONNAIRE ITEMS USED TO MEASURE
GENERAL HEALTH PERCEPTIONS

Items[a]	Investigator(s)
In general (all in all) would you say your own health is excellent, good, fair, or poor?	Madow (1969) Renne (1970) Berkman (1971) Renne (1971) Cole and Lejeune (1972) Renne (1974) Spreitzer and Snyder (1974)
How would you characterize your health overall?	Reynolds, Rushing, and Miles (1974)
How is your health in general?	Schnore and Cowhig (1959-60) Gaitz and Scott (1972)
Compared with others your age, would you say your health was better (the same) or worse?	DiCicco and Apple (1958) Henley and Davis (1967) Fabrega and Roberts (1972) Renne (1974) Warren (1974)
How would you rate your health (at the present time) (these days) right now?	Suchman, Phillips, and Streib (1958) Thompson and Streib (1958) DiCicco and Apple (1958) Maddox (1963) MacPherson (1972) Palmore and Liukart (1972) Tissue (1972) Cutler (1973) Heyman and Jeffers (1973) Warren (1974)
Would you say that your health is good, average, or poor?	Edwards and Klemmack (1973)
To what degree are you satisfied with your health status?	Tornstam (1975)
Has your health changed during the past year?	Thompson and Streib (1958)
Is your health better, worse, about the same as it was 10 years ago?	DiCicco and Apple (1958)
Considering any sickness or injury you might have, check the item that best describes you	Bergner, Bobbitt, Pollard, et al. (1976)

[a]See Table 5 for response choices.

Table 3

EXAMPLES OF LABELS USED TO DESCRIBE GENERAL HEALTH RATINGS

Labels	Investigator(s)
Self-evaluation of (physical, general health)	Thompson and Streib (1958) Madow (1967) Hochstim, Athanasopoulos, and Larkins (1968) Shanas, Townsend, Weddarburn, et al. (1968) Berkman (1971) Bergner, Bobbitt, Pollard, et al. (1976)
Self-assessed health (status)	Maddox (1964) Spreitzer and Snyder (1974) Fine (1975)
Self-health ratings	Suchman, Phillips, and Streib (1958) Friedsam and Martin (1963) Heyman and Jeffers (1963) Leveson (1972) Palmore and Luikart (1972) Tissue (1972) Nowlin (1974)
Self-health evaluation	Osborn (1973)
Self-report of physical health	Gaitz and Scott (1972)
Self-concept of oneself as healthy	Kessel and Shepard (1965)
Perceived (self-perceptions of) health status	Suchman, Phillips, and Streib (1958) Schnore and Cowhig (1959-60) Fabrega and Roberts (1972) Edwards and Klemmack (1973) Garrity (1973) Reynolds, Rushing, and Miles (1974) Warren (1974) Tornstam (1975)
Subjective (assessment or estimate of) health	Maddox (1964) Desroches, Kaiman, and Ballard (1967) Henley and Davis (1967) Cole and Lejeune (1972) Crawford and Hooper (1973) Cutler (1975)

Table 3 (continued)

Labels	Investigator(s)
Satisfaction with health	DiCicco and Apple (1958) Renne (1970, 1971, 1974)
General health rating	Suchman, Phillips, and Streib (1958) Renne (1970)
Level of health	Pratt (1971)
Poor health	Ludwig and Gibson (1969)
Current health	Desroches, Kaiman, and Ballard (1967)
Past health	Desroches, Kaiman, and Ballard (1967)
Lifetime health	Pratt (1972)

USES OF GENERAL HEALTH RATINGS

As illustrated in Table 1, general health ratings have been used to (a) study relationships among health constructs (23 studies), (b) explain health and illness behavior (7 studies), (c) describe the health of populations (7 studies), and (d) clarify measurement issues (2 studies). Most study designs were cross-sectional, regardless of whether general health ratings were treated as independent or dependent variables in the analysis.

ITEM SCORING METHODS

Table 4 summarizes response choices offered for general health ratings. Most studies based on single-item general health measures used from two to ten response categories. For analytic purposes, five investigators dichotomized responses into favorable/unfavorable categories; one trichotomized responses; and others concentrated on the fair, poor, or worse categories. Most therefore emphasized the negative end of the continuum of general health perceptions.

Table 5 summarizes the methods used to score general health ratings. Investigators rarely stated whether the level of measurement was nominal, ordinal, interval, or ratio. In nearly all instances, levels of measurement reported in Table 5 were inferred from authors' descriptions and from statistical methods. General health ratings were usually scored as nominal or interval.

Because statistics that treat data as nominal ignore information about the rank order of responses, investigators who used them may have unnecessarily lost statistical power in their analysis of general health perceptions data. General health rating scales probably yield at least ranked data. In the absence of empirical studies supporting the interval scale assumptions underlying the analytic techniques used by other investigators, however, their appropriateness cannot be evaluated.

Table 4

EXAMPLES OF RESPONSE CATEGORIES FOR GENERAL HEALTH RATING ITEMS

Response Categories	Investigator(s)
Excellent, good, fair, poor	Schnore and Cowhig (1959-60) Maddox (1963) Renne (1970, 1971) Berkman (1971) Cole and Lejeune (1972) Palmore and Luikart (1972) Spreitzer and Snyder (1974)
Fair or poor	Hochstim, Athanasopoulos, and Larkins (1968)
Poor or worse	Ludwig and Gibson (1969)
Excellent, good, fair, poor, very poor	Suchman, Phillips, and Streib (1958) Friedsam and Martin (1963) Leveson (1972) Cutler (1973)
Excellent, good, fair, not very good, poor	Reynolds, Rushing, and Miles (1974)
Good, average (fair), poor	Shanas, Townsend, Weddarburn, et al. (1968) Tissue (1972) Edwards and Klemmack (1973) Osborn (1973) Warren (1974)
Good, pretty good, not so good	MacPherson (1972)
Worse, the same, better, much better	Fabrega and Roberts (1972)
Very good, pretty good, not so good, poor	Gaitz and Scott (1972)
Very poor, poor, fair for my age, fair, good, good for my age, excellent, excellent for my age	Heyman and Jeffers (1963) Fine (1975)
Very good, good, fair, poor (very poor)	Henley and Davis (1967)
Better, same, worse	Warren (1974)
Not sick, very mildly sick, mildly sick, moderately sick, quite sick, very sick, extremely sick	Bergner, Bobbitt, Pollard, et al. (1976)

Table 5

SUMMARY OF RESPONSE CHOICES, SCORING, AND LEVEL OF
MEASUREMENT FOR GENERAL HEALTH RATINGS
REPORTED IN THE LITERATURE

Investigator(s)	Response Choices/Scoring[a]	Level of Measurement[b]
DiCicco and Apple (1958)	Five-item index of satisfaction with health (response categories not reported)	Interval
Suchman, Phillips, and Streib (1958)	Dichotomized: Favorable = excellent, good; Unfavorable = fair, poor, very poor	Nominal
Thompson and Streib (1958)	Five-point Guttman scale One extreme: those who rate health as good or excellent, deny having a physical problem, say health has not changed, have not seen doctor during previous year	Ordinal
Schnore and Cowhig (1959-1960)	Excellent, good, fair, poor	Nominal
Friedsam and Martin (1963)	Dichtomized from: (1) very poor; (2) poor; (3) fair; (4) good; (5) excellent[c]	Interval
Heyman and Jeffers (1963)	Dichotomized from: very poor, poor, fair for my age, fair, good, good for my age, excellent, excellent for my age[c]	Nominal
Maddox (1963, 1964)	Dichotomized from: excellent, good vs. fair, poor	Nominal
Kessel and Shepard (1965)	Not given	Nominal
Desroches, Kaiman, Ballard (1967)	Five-point scale (5 = worst rating)	Interval
Henley and Davis (1967)	Dichotomized from: very good, good vs. fair, poor	Nominal
Madow (1967)	Excellent, good, fair, poor	Nominal

Table 5 (continued)

Investigator(s)	Response Choices/Scoring[a]	Level of Measurement[b]
Hochstim, Athanasopoulos, and Larkins (1968)	Considered only negative self-evaluations of health: fair and poor	Interval
Shanas, Townsend, Weddarburn, et al. (1968)	Good, fair, poor	Nominal
Ludwig and Gibson (1969)	Considered only negative self-evaluations of health: poor and worse	Nominal
Renne (1970)	Excellent, good, fair, poor	Nominal
Berkman (1971)	Excellent, good, fair, poor, combined with number of hospital days during the past five years; response choices then dichotomized[c]	Nominal
Renne (1971)	Excellent, good, fair, poor	Interval
Cole and Lejeune (1972)	Excellent, good, fair, poor (analyses based on percentages who responded fair or poor)	Interval
Fabrega and Roberts (1972)	(1) Worse, (2) the same, (3) better, (4) much better	Interval
Gaitz and Scott (1972)	Very good, pretty good, not so good, poor	Interval
Leveson (1972)	Excellent or good, fair or unknown, poor or very poor	Nominal
Palmore and Luikart (1972)	Ten-point scale: 9 = perfect health to 0 = most serious illness	Interval
Pratt (1972)	Not given	Interval
Tissue (1972)	General health: good, fair, poor Health compared with health of peers: healthier/same, sicker	Ordinal

Table 5 (continued)

Investigator(s)	Response Choices/Scoring[a]	Level of Measurement[b]
Crawford and Hooper (1973)	Better, the same, worse (than other women same age)	Nominal
Cutler (1973)	Excellent, good, fair, poor	Interval
Edwards and Klemmack (1973)	Good, average, poor	Interval
Garrity (1973)	Ten-point scale	Interval
Osborn (1973)	Good, fair, poor	Nominal
Nowlin (1974)	Six-point scale: 5 = excellent health to 0 = very poor health	Interval
Renne (1974)	Four-point scale combined response to question regarding health in relation to that of others and general health rating: (1) better than others, excellent (2) better and good or same as others and excellent (3) same and good (4) either worse than others or fair or poor	Interval
Reynolds, Rushing, and Miles (1974)	(1) Poor; (2) not very good; (3) fair; (4) good; (5) excellent	Interval
Spreitzer and Snyder (1974)	Excellent, good, fair, poor	Interval
Warren (1974)	General health: good, fair, poor Health compared with that of others same age; better, same, worse	Nominal
Fine (1975)	Five-point scale ranging from very poor to very good	Interval

Table 5 (continued)

Investigator(s)	Response Choices/Scoring[a]	Level of Measurement[b]
Tornstam (1975)	Five response choices (not reported) regarding satisfaction with health status	Interval
Bergner, Bobbitt, Pollard, et al. (1976)	Self-assessment of sickness: not sick, very mildly sick, mildly sick, moderately sick, quite sick, very sick, extremely sick Self-assessment of dysfunction: not at all, very slightly, slightly, moderately, greatly, very greatly, extremely	Interval

[a]Numerical scoring included in table if reported in article.

[b]Reported in article or inferred from statistical techniques.

[c]Cutting point for dichotomous scoring not reported.

GENERAL HEALTH RATING SCALES

The literature review revealed only seven examples of scales constructed from general health rating items.[2] Table 6 summarizes information about these scales, including scaling methods (where reported), content of items, and details of scoring methods. Both the Method of Summated Ratings (Likert, 1932) and Scalogram Analysis (Guttman, 1944) have been used. Four scales (DiCicco and Apple, 1958; Thompson and Streib, 1958; Berkman, 1971; Pratt, 1971) included an item regarding use of medical care services (either physician visits or hospitalization). Because responses to such items reflect, among other things, accessibility of medical care services and willingness to adopt the patient role, the appropriateness of their inclusion in a scale to measure general health perceptions is questionable on grounds of validity.

RELIABILITY AND STABILITY[3]

No investigators who used single-item measures of health perceptions reported reliability estimates. Two reported reproducibility coefficients for scales constructed according to the criteria of Scalogram Analysis; when high, these coefficients can be taken as evidence of both reproducibility and reliability. Both scales included multiple general health rating items along with an item to measure use of medical care services. Thompson and Streib (1958) reported high reproducibility coefficients

[2] Since completion of the literature review, Becker and his colleagues (Maiman, Becker, Kirscht, et al., 1977) have published information about the scaling of variables defined by the Health Belief Model. Included were measures that fit the definition of general health perceptions used in the HIS (e.g., perceived severity of illness, susceptibility).

[3] See Vol. I, Sec. II, which discusses HIS methods and standards for evaluating reliability and stability.

Table 6

Summary of Published Information Regarding General Health Rating Scales

Investigator	Suchman et al. (1958)	Thompson and Streib (1958)[a]	DiCicco and Apple (1958)[b]	Pratt (1971)
Scaling Method Used	Not reported	Scalogram Analysis	Scalogram Analysis	Not reported
Scale Name	--	--	Index of Health	Level of Health Index
Items/Response Categories	Self-Ratings a. How would you rate your health at the present time? (very poor, poor, fair, good, excellent) b. Do you have any particular physical or health problems at present? (yes, no) Physician Rating c. How would you rate the subject's health at the present time? (very poor, poor, fair, good, excellent)	a. Has your health changed during the past year? (yes, no) b. How would you rate your health at the present time? (excellent, good, fair, poor) c. Do you have any particular physical or health problems at present? (yes, no) d. Have you been seen by a doctor during the past year? (yes, no)	Item A. Number of days in bed during previous year[c] Item B. Index of Satisfaction with Health a. How would you say your health is now? b. Do you think your health is better or worse than that of other people your age? c. Is your health better or worse or about the same as it was 10 years ago? d. Does it keep you from doing things? e. Does it keep you from seeing people? Item C. Whether any present trouble with health[c]	a. Respondent's rating of present level of health[c] b. Occurrence of illness during preceding 2 weeks[c]
Scoring Method	Responses to first and third items dichotomized as favorable (responses of excellent or good) vs. unfavorable (fair, poor, very poor); three dichotomized scores combined	Five scale types ranging from extremes of no changes/excellent or good/no problems/no care at positive end to change for worse/fair or poor/problems/medical care at negative end of scale.	Four scale types ranging from most to least healthy; response categories and scoring details not reported	Not reported (high score indicative of good health)

Table 6 (continued)

	Berkman (1971)	MacPherson (1972)	Renne (1974)
Investigator			
Scaling Method Used	Not reported	Summated Ratings[d]	(See scoring method)
Scale Name	Adult Physical Health	--	--
Items/Response Categories	a. All in all, would you say your health is generally excellent, good, fair, or poor? b. During the past five years have you ever stayed overnight at a hospital, sanitarium, or nursing home? (Do not count childbirth) (yes, no) If yes, how many separate times did you stay there?	a. How would you say your health is these days? b. Taking all things together, how would you say things are these days? c. In general, how satisfying do you find the way you are spending these days?	a. Is your health better or worse than the health of most other people your age? b. All in all, would you say your health is generally excellent, good, fair, or poor?
Scoring Method	Responses to items dichotomized such that 1/4 to 1/3 of sample fell into "stressful" category; two dichotomized scores combined	Response categories assigned values from 1 to 3; item scores summed to score scale	Responses combined to form 4-point subjective health index: 1) Better than others/excellent 2) Better than others/good or same as others/excellent 3) Same as others/good 4) Worse than others/fair or poor Scores converted to ridit values to express the probability that a member of a subgroup is worse off than someone in a specified distribution

[a] Response categories inferred from article; coefficient of reproducibility = 0.95 to 0.96 on three administrations.

[b] Coefficient of reproducibility > 0.90.

[c] Actual item wording not reported in article.

[d] Method of scaling inferred from article.

(0.95 to 0.96) for their four-item scale. DiCicco and Apple (1958) reported a reproducibility coefficient greater than 0.90 for their seven-item Index of Health, which included a five-item Index of Satisfaction with Health subscale (coefficient not reported for the subscale; see Table 6).

Four studies provided data on the stability (intertemporal association) of general health ratings. Suchman, Phillips, and Streib (1958) and Heyman and Jeffers (1963) reported cross-tabulations of general health rating scores obtained from the same respondents at two times approximately two and three years apart, respectively. Although they did not report coefficients of association, phi coefficients could be computed from their published data. Phi coefficients of 0.07 and 0.43 were computed for data reported by Suchman, Phillips, and Streib (1958) and Heyman and Jeffers (1963), respectively; both coefficients were statistically significant (p < 0.05). Results from a study in progress (Rushing, personal communication, 1977) included phi coefficients for dichotomously scored general health ratings. Coefficients ranged from 0.57 to 0.52 for intervals of one to three years between administrations; stability decreased as the interval increased. Maddox (1963) reported no change in self-assessments of health status for 76 percent of elderly respondents (ages 60 to 94, median of 70 years) interviewed approximately three years apart; the stability coefficient was significant at p < 0.001.

In summary, only two of the 39 articles discussed reliability of general health ratings scores, three explored the stability of these ratings over time, and one study in progress provided stability data. Although available information is therefore sparse, it suggests that ratings are reliable and reproducible and that their stability exceeds chance expectations.

VALIDITY OF GENERAL HEALTH RATINGS

In most studies, the validity of general health ratings (the extent to which they measure what they were intended to measure) has apparently been presumed from the content of questionnaire items. Although many published articles did not explicitly address validity issues, some reported results that were useful for that purpose. Information regarding associations between general health ratings and the following types of variables are relevant in assessing validity: (1) general health perceptions defined or measured differently, (2) measures of specific health status components (physical, mental, and social), (3) health and illness behavior (e.g., use of medical care services), and (4) age. If general health ratings are valid, the following relationships would be hypothesized: (1) strong associations among general health ratings based on different methods, (2) substantial associations between general health ratings and measures of specific health components, (3) positive associations between ratings by respondents and their physicians, (4) negative associations between ratings and illness behavior (e.g., recent use of services), and (5) less favorable general health ratings with increasing age.

These hypotheses were based on measurement theory and on theories about the origin and effects of general health perceptions. The first hypothesis follows from an approach to measurement validation based on study of convergent and discriminant validity using a multitrait-multimethod matrix (see Campbell and Fiske, 1959); measures of the same construct (general health perceptions) using

different methods should be significantly correlated. The second hypothesis is based on the assumption that general health ratings assess a factor theoretically common to measures of physical, mental, and social components of health status. Thus, general health ratings should be significantly related to all measures of these components. The second hypothesis should also hold true because information about changes in specific components of health status (e.g., physical ability, level of depression) should be taken into account at the time of rating health in general.

Because physicians and their patients consider the same information (at least in part) when rating the patient's health, significant correlations between self- and physician-ratings of health status were hypothesized. Patients also tend to rate their health consistently with what physicians tell them; hence, the third hypothesis. The fourth hypothesis was based on the assumption that medical care use is influenced by need, among other factors. One way of defining need is in terms of a personal assessment of health. According to theory, need (perceived health) directly predicts use of services. Finally, measures of physical health in terms of functioning yield less favorable scores with increasing age (see Vol. II). Thus, negative associations between measures of general health perceptions (favorably scored) and age were hypothesized. Because all studies that used general health ratings were based on cross-sectional data, studies of these relationships did not answer questions regarding the predictive validity of the ratings, or the nature of the cause and effect relationships between ratings and other measures of health or use of services.

Table 7 summarizes validity findings reported in the literature. Data sources for validity variables included self-report (S), physician assessment (P), and interviewer assessment (I). To facilitate comparison of results across studies, signs associated with coefficients in Table 7 were adjusted to indicate the direction of relationships that would have been observed if high numbers had always been assigned to favorable health ratings and low numbers to unfavorable. Similarly, signs were adjusted to indicate relationships that would have been observed had high scores been assigned to other variables in a manner consistent with their names (e.g., a high score on "life satisfaction" indicating greater satisfaction).

Associations Among General Health Ratings

Very few investigators reported associations among general health ratings based on different methods or definitions. Absence of such data undoubtedly indicates that the majority used one item to measure general health and, because their studies were not designed to explore the psychometric properties of their measures, did not include other rating items that would hve been useful in studying these properties.

In studies that reported associations among different measures of general health perceptions, general health was usually defined as a personal assessment of health (in a manner consistent with the definition offered in the Introduction) and has been correlated with measures of attitude toward personal health, level of health necessary to achieve personal satisfaction, and amount of worry about health (see Table 7). Heyman and Jeffers (1963) reported a significant positive association between their single-item general health rating and an eight-item mea-

Table 7

SUMMARY OF PUBLISHED FINDINGS REGARDING ASSOCIATIONS BETWEEN GENERAL
HEALTH RATINGS AND HEALTH-RELATED VARIABLES

| Health and Health-Related Variables | Data Source | Association[a] | | | Investigator(s) |
		Sign	Statistic	Method	
Other General Health Ratings					
Health attitude	S	+	13.2***	χ^2	Heyman and Jeffers (1963)
	S	+	27.5	χ^2	Heyman and Jeffers (1963)
Health level necessary for personal satisfaction	S	+	.25	r	Tornstam (1975)
Health worry	S	–	6.2*	χ^2	Friedsam and Martin (1963)
	S	–	NG**	χ^2	Suchman, Phillips, and Streib (1958)
	S	–	–.73***	γ	Tissue (1972)
Functions/Dysfunctions					
Function Status Index	S	+	.59***[b]/.72***[c]	r	Reynolds, Rushing, and Miles (1974)
	S	+	.66***	γ	Tissue (1972)
Sickness Impact Profile	S	+	.54***	r	Bergner, Bobbitt, Pollard, et al. (1976)
Expectation of not working	S	–	–.22***[d]/ –.27***[e]	r	Cole and Lejeune (1972)
Self-defined role performance	S	+	NG	%	Cole and Lejeune (1972)
Bed days	S	–	–9.1**	χ^2	Friedsam and Martin (1963)
	S	–	NG**	%	Suchman, Phillips, and Streib (1958)
Incapacity	S	–	NG	%	Shanas, Townsend, Weddarburn, et al. (1968)
Work days prior to heart attack	S	+	.29*	r	Garrity (1973)
Return to work after heart attack	S	+	.59**	r	Garrity (1973)
Mortality					
Deceased	S	–	18.03***	χ^2	Heyman and Jeffers (1963)
Physical Health					
Number of chronic conditions	S	–	NG	%	Madow (1967)
	S	–	172.6***	χ^2	Osborn (1973)
Presence of chronic conditions	S	–	NG	%	Renne (1971)
Number of symptoms	S	–	–.41 (current health)	r	Desroches, Kaiman, and Ballard (1967)
	S	–	–.54** (past health)	r	Desroches, Kaiman, and Ballard (1967)
	S	–	–.44	path r	Tornstam (1975)
Presence of:					
aches	S	–	–.46***	r	Tornstam (1975)
serious disease	S	–	–.48***	r	Tornstam (1975)
visible impairment	S	–	–.40***	r	Tornstam (1975)
general exhaustedness	S	–	–.38***	r	Tornstam (1975)
vascular disease	S	–	–.42***	r	Tornstam (1975)
sensory impairment	S	–	NG	%	Shanas, Townsend, Weddarburn, et al. (1968)
Number of health problems	S	–	–.50***	γ	Tissue (1972)
Number of illnesses	S	–	NG	%	Cole and Lejeune (1972)
Mobility	S	+	.49***	r	Tornstam (1975)
	S	+	NG	%	Shanas, Townsend, Weddarburn, et al. (1968)
Zeman's Scale of Activity	I	+	.54*	r	DiCicco and Apple (1958)
Ability to go outdoors without difficulty	S	+	.37***	r	Fine (1975)
Ability to go up and down stairs	S	+	.40***	r	Fine (1975)
Bodily concern	S	–	NG	%	Maddox (1964)
Self-rating of physical health	S	+	NG	ridits	Renne (1974)

Table 7 (continued)

Health and Health-Related Variables	Data Source	Sign	Association[a] Statistic	Method	Investigator(s)
Mental Health					
Index of psychological well-being	S	+	NG	ridits	Berkman (1971)
Life satisfaction	S	+	.33***	η	Cutler (1973)
	S	+	.24	path r	Tornstam (1975)
	S	+	.19*	r	Edwards and Klemmack (1973)
	S	+	NG*	x^2	Henley and Davis (1967)
	S	+	.25**	r	Spreitzer and Snyder (1974)
	S	+	NG	%	Thompson and Streib (1958)
Happiness	S	+	5.7*	x^2	Friedsam and Martin (1963)
	S	+	.34**	r	Gaitz and Scott (1972)
	S	+	NG**	%	Suchman, Phillips, and Streib (1958)
	S	+	.17	γ	Tissue (1972)
Self-conception as old	S	–	2.0	x^2	Friedsam and Martin (1963)
Anomie scale	S	–	8.2**	x^2	Gaitz and Scott (1972)
Loneliness	S	–	NG	%	Shanas, Townsend, Weddarburn, et al. (1968)
Usefulness feelings	S	+	7.8**	x^2	Friedsam and Martin (1963)
Feel blue	S	–	5.0*	x^2	Friedsam and Martin (1963)
Dejection	S	–	NG	%	Thompson and Streib (1958)
Hopelessness	S	–	NG	%	Thompson and Streib (1958)
Langner's 22-item score	S	–	-.47**	r	Gaitz and Scott (1972)
Self-satisfaction	S	+	.17**	r	Gaitz and Scott (1972)
Morale	S	+	NG	%	Maddox (1964)
Marital satisfaction index	S	+	NG	%	Renne (1971)
Marital organization					
marital power	S	+	NG[f]	t	Pratt (1972)
sex role differentiation	S	+	NG[g]	t	Pratt (1972)
companionship	S	+	NG*	t	Pratt (1972)
Perceived self-value	S	+	.19	path r	Tornstam (1975)
Social Health					
Voluntary association	S	+	.24[b]	η	Cutler (1973)
Social health index	S	+	NG	ridits	Renne (1974)
Perceived social value	S	+	.14	path r	Tornstam (1975)
Degree of isolationism	S	–	NG	path r	Tornstam (1975)
Degree of interaction	S	+	.19	r	Tornstam (1975)
Health of Others					
Health compared with that of peers	S	+	.77***	γ	Tissue (1972)
	S	+	NG	%	Warren (1974)
	S	+	NG	%	Shanas, Townsend, Weddarburn, et al. (1968)
Physician Assessment					
Physician's general rating	P	+	5.0*	x^2	Friedsam and Martin (1963)
	P	+	NG**	x^2	Suchman, Phillips, and Streib (1958)
Presence of chronic condition	P	–	.08	r	Garrity (1973)
Physician's ratings of function	P	+	2.23*[b] 17.85***[c]	x^2	Heyman and Jeffers (1963)
	P	+	NG	%	Maddox (1964)
	S	+	NG	ridits	Renne (1974)
	P	+	.43* (current health)	r	Desroches, Kaiman, and Ballard (1967)
	P	–	-.13 (past health)	r	Desroches, Kaiman, and Ballard (1967)
	P	–	-.08	r	Bergner, Bobbitt, Pollard, et al. (1976)
Depression	P	–	NG*	F	Nowlin (1974)

Table 7 (continued)

Health and Health-Related Variables	Data Source	Association[a]			Investigator(s)
		Sign	Statistic	Method	
Health and Illness Behavior					
Frequency of medical care	S	–	NG*	t	Fabrega and Roberts (1972)
Appropriateness of use of services	S	+	NG*	t	Pratt (1971)
Participation in health plan					
Total population	S	–	NG**	χ^2	Leveson (1972)
Heart disease subgroup	S	–	NG	χ^2	Leveson (1972)
Number of doctor visits	S	NG	4.9*	χ^2	Friedsam and Martin (1963)
Last contact with doctor	S	–	-.36**	γ	Tissue (1972)
Regular checkups	S	–	-.09	γ	Tissue (1972)
Last time in hospital	S	–	-.31*	γ	Tissue (1972)
Personal health practice	S	+	NG	NG	Pratt (1971)
Knowledge of Health Matters					
Health knowedge	S	+	NG*	t	Pratt (1971)

[a]Abbreviations used in these columns:

      NG = not given

      * $p < 0.05$

      ** $p < 0.01$

      *** $p < 0.001$

      χ^2 = Chi-square value

      r = product-moment correlation coefficient

      γ = gamma correlation coefficient

      % = differences in proportions of population at different levels

  ridit = "relative to an identified population"; an average ridit for a given subgroup expresses the probability that a member of the subgroup is "worse off" than one in the identified population

  path r = coefficient from a path analysis

      η = eta correlation coefficient

      F = F-test value

      t = t-test value

[b]For total population.

[c]For sample population.

[d]New York Survey.

[e]Camden survey.

[f]Lifetime health significant at $p < 0.05$ for wives only.

[g]Lifetime and present health significant at $p < 0.05$ and $p < 0.005$ for wives only.

sure of general attitude toward health.[4] General health ratings were significantly and negatively associated with measures of health worry in all studies of this relationship (Suchman, Phillips, and Streib, 1958; Friedsam and Martin, 1963; Tissue, 1972); Tissue (1972) found the association between these two perceptions was very strong. Two associations among general perceptions were not statistically significant; however, the direction of the relationships was consistent with validity hypotheses (see Table 7).

Function/Dysfunction

Several studies compared measures of functions and dysfunctions (function status), which probably reflect variation in physical, mental, and social components of health status, with general health ratings (Suchman, Phillips, and Streib, 1958; Shanas, Townsend, Weddarburn, et al., 1968; Cole and Lejeune, 1972; Tissue, 1972; Garrity, 1973; Reynolds, Rushing, and Miles, 1974; Bergner, Bobbitt, Pollard, et al., 1976). Associations between favorably defined function status measures and general health ratings were significant and positive in the five studies in which significance was reported or could be determined from reported data. Associations between general health ratings and negatively defined function status variables were negative, as hypothesized, and significant in all studies that reported significance tests. Without reliability estimates, the possibility that measurement error may have led to underestimation of some of these associations cannot be dismissed. Nonetheless, several substantial associations were reported, suggesting that general health ratings are valid in relation to differences in function status.

Mortality

Heyman and Jeffers (1963) reported a significant association between general health ratings and mortality in a longitudinal study. Comparisons between scores for persons still alive and those who had died before a second administration indicated that those who had died were more likely than others to have rated their health unfavorably on the first administration.

Physical Health

Measures of physical health compared with general health ratings included self-assessments of physical ability, self-reported chronic conditions and symptoms indicating physical health, and personal assessments of physical health status. In general, nearly all published associations between general health ratings and measures of physical health were in the hypothesized direction. The magnitude of the majority of associations, which were likely to be lower-bound estimates because of lack of perfect reliability, indicated substantial relationships and supported the validity of general health ratings in relation to physical health status.

People who rated their general health as fair or poor also reported a greater number of chronic conditions (Madow, 1967; Garrity, 1973; Osborn, 1973); only in the last study was the relationship significant. Renne (1971) found that people

---

[4] The eight-item health attitude measure was constructed by Cavan, Burgess, Havighurst, et al. (1949).

reporting disability and chronic conditions were more likely to rate their health unfavorably than were those who reported minor or no symptoms.

Significant associations between number and presence of health problems or symptoms and general health ratings were reported by Desroches, Kaiman, and Ballard (1967); Cole and Lejeune (1972); Tissue (1972); and Tornstam (1975). People who reported more health problems or symptoms were more likely to rate their health unfavorably (Desroches, Kaiman, and Ballard, 1967). This relationship was significant for self-ratings of previous health, but not for self-ratings of current health. Cole and Lejeune also found that the number of illnesses reported was inversely related to favorable health ratings, although significance tests were not reported. In addition to showing that a substantial proportion of the variance in subjective health ratings and number of symptoms was shared, Tornstam reported relationships between symptom groups and general health ratings. General health ratings were inversely related to presence of aches, serious disease, visible impairment, general exhaustion, and vascular disease.

General health ratings were less favorable for respondents with higher degrees of sensory impairment and immobility (Shanas, Townsend, Weddarburn, et al., 1968). The latter association was confirmed by Tornstam (1975).

Subjective ratings of bodily concern and physical health were also associated with general health ratings (as hypothesized) in studies reported by Maddox (1964) and Renne (1974). Individuals with more bodily concern and less favorable perceptions of physical health tended to evaluate their general health less favorably; significance tests were not reported.

Mental Health

Several studies included measures of mental health constructs and general health ratings. In general, associations tended to be positive with positively defined mental health constructs (e.g., life satisfaction) and negative with those negatively defined (e.g., depression).

Six studies reported positive associations between general health ratings and life satisfaction. This relationship was statistically significant in four (Cutler, 1973; Edwards and Klemmack, 1973; Henley and Davis, 1967; Spreitzer and Snyder, 1974) and nonsignificant in one (Tornstam, 1975); another study (Thompson and Streib, 1958) did not test for significance. General health ratings and "happiness" were significantly and positively related in studies reported by Suchman, Phillips, and Streib (1958), Friedsam and Martin (1963), and Gaitz and Scott (1972); Gaitz and Scott also reported a significant positive relationship between ratings of satisfaction with self and general health. General health ratings were also positively associated with morale (Maddox, 1964), measured by the Havighurst scale (Maddox and Eistorfer, 1962). The strength and significance of the association were not reported.

Berkman (1971) studied an eight-item index of psychological well-being labeled "mental health risk" and observed unfavorable health ratings in conjunction with increases in mental health risk.

Friedsam and Martin (1963) reported less favorable general health ratings when respondents felt anomic or "blue." Shanas, Townsend, Weddarburn, et al. (1968) correlated general health ratings with measures of loneliness; general health ratings were less favorable when respondents felt lonely. Measures of dejection and

hopelessness also tended to be inversely related to general health ratings (Thompson and Streib, 1958). The higher the general health rating, the less the dejection and hopelessness.

Gaitz and Scott (1962) reported a negative correlation between a general health rating and impaired life functioning due to psychiatric symptoms, defined using the 22-item Langner scale.

Renne (1971) found that satisfaction with marriage was related to general health ratings. Divorced and unhappily married men did not differ in general health ratings; unhappily married women tended to rate their health less favorably than divorced women of the same age. Significance tests were not reported. Pratt (1972) reported that the higher the level of husband-wife companionship, the more favorable the general health ratings (both lifetime and present). The more egalitarian the marital relationship and the less role differentiation, the more likely wives were to report high levels of general health.

The pattern of associations supported the validity of general health ratings in relation to mental health status. The strength of the relationships between these constructs could not, however, be estimated from available data. Coefficients of association were reported less often than for physical health and other validity variables, and information regarding measurement error was rarely reported.

Social Health

Associations between social health constructs and general health ratings were identified in only four publications. In general, more favorable ratings were related to positive aspects of social health. The strength of these relationships was usually less than that observed between general health ratings and other components of health status.

Cutler (1973) reported that membership in voluntary organizations among the aged was significantly related to general health: Those who rated their health favorably had significantly higher scores on an index of voluntary participation in organizations. Renne (1974) used data on employability, family relationships, community involvement, and personal contacts outside the family to develop the Social Health Index, and compared index scores with general health ratings. Favorable ratings were associated with positive assessments of one's marriage, qualifications for employment, and high levels of involvement with people outside the family; significance tests were not reported. Tornstam (1975) used path analysis to study the relationship between satisfaction with health status and perceived "social value." The relationship was positive, indicating that (with other variables held constant) perceived social value increased with greater satisfaction with health status. Furthermore, general health ratings were positively associated with social interaction and negatively associated with isolation.

Health of Others

Tissue (1972) observed a very strong association between self-ratings of personal health and ratings of one's health in relation to other people in the respondent's age group (suggesting that these two approaches may measure the same construct). Similar findings were reported by Shanas, Townsend, Weddarburn, et

al. (1968) and Warren (1974); those who perceived their health as favorable were more likely to report that their health was better than that of others their age.

Physician Assessments

Of the nine studies that compared general health ratings with physician assessments of health status, six reported significant associations. All were in the hypothesized direction. Physician assessments of functional activity level tended to be higher for those who rated their current health favorably (Desroches, Kaiman, and Ballard, 1967). Heyman and Jeffers (1963) reported significant and positive relationships between physician assessments and self-ratings at two times three years apart. Although significance tests were not reported, trends in two studies (Maddox, 1964; Renne, 1974) also suggested that general health ratings were closely related to physician assessments of physical health. Bergner, Bobbitt, Pollard, et al. (1976), however, reported that self-assessments of sickness by speech pathology patients did not correlate significantly with assessments made by speech pathologists. As hypothesized, Nowlin (1974) reported a significant negative relationship between self-ratings of general health and physicians' assessments of the respondents' depression.

When rating health in general, physicians in two studies (Suchman, Phillips, and Streib, 1958; Friedsam and Martin, 1963) tended to agree with their patients. In both studies, associations were significant and positive.

Health and Illness Behavior

Of the nine comparisons of general health ratings and variables defining health and illness behavior, seven indicated significant associations in the hypothesized direction. Favorable health ratings were significantly associated with better use of professional services (Pratt, 1972); no contact with a doctor in the past six months (Tissue, 1972); no contact with a doctor in the past year (Fabrega and Roberts, 1972); no admission to the hospital in the past five years (Tissue, 1972); less contact with a doctor or fewer hospital admissions in the past year (Suchman, Phillips, Streib, et al., 1958); less participation in a health maintenance program (Leveson, 1972); and good personal health practices (Pratt, 1971). In general, those who had used services less often were more likely to rate their health favorably. When use of services was measured, appropriate patterns of use were associated with favorable general health ratings (Pratt, 1971). Regularity of checkups was not related to general health ratings in the study reported by Tissue (1972). Leveson (1972) found no relationship between health ratings and participation in a health maintenance program for people with chronic heart conditions.

Pratt (1971) found no relationship between level of health knowledge (adequacy of respondent's information concerning common health matters) and general health ratings. Pratt also reported a significant relationship between personal health practices and general health perceptions. The higher the quality of personal health practices (defined from responses to questions dealing with sleep, exercise, smoking, and nutrition), the higher the health rating.

Published associations between general health ratings and health and illness behavior were thus consistent with the hypothesized predictive validity of ratings. The cross-sectional nature of most data, however, precluded clarification of the

exact nature of the cause and effect relationships involved. General health perceptions may be brought in line with behavior after the behavior occurs (Bem, 1972).

Age Differences in General Health Ratings

Associations between age and general health ratings were negative, as hypothesized, in four of seven studies (see Table 8). With increasing age, welfare mothers were more likely to give poor general health as a reason for preferring not to work (Cole and Lejeune, 1972). This association was also reported for women who had not been sick during the previous year; thus, older respondents (who by their own account had not been sick) were more likely than younger respondents to give poor health as a reason not to work. Gaitz and Scott (1972) found that middle-aged and older respondents tended to rate their health less favorably than did younger people. When sex and skill level were controlled for, the linear positive relationship between age and reports of poor health remained. Similarly, Palmore and Luikart (1972) and Schnore and Cowhig (1959-60) found negative relationships between general health ratings and age.

Maddox (1964) reported that older people between the ages of 60 and 94 tended to be more optimistic about their health than younger people within that age range. In other studies of older populations, Heyman and Jeffers (1963) and Desroches, Kaiman, and Ballard (1967) found no associations between self-health ratings and age.

Conclusions

Without reliability data, the precise strength of reported associations among general health ratings, or between general health ratings and other health measures, health and illness behavior, and age could not be estimated. Insignificant associations may, in some instances, have been due to measurement error. Despite dissimilarities in measurement methods and populations studied, the weight of published evidence is consistent with the hypothesis that general health ratings are valid measures of health status. In many instances, these associations were substantial. Published validity coefficients are probably lower-bound estimates of true relationships among these variables (because of lack of perfect reliability, particularly for single-item measures).

The direction of associations among general health ratings, and between general health ratings and other health and health-related constructs, was nearly always consistent with what would be predicted if general health ratings actually reflect an underlying health construct. Associations among general health ratings scored in the same direction were positive; correlations with positive definitions of physical, mental, and social health were consistently positive; and correlations with negative definitions of health status were consistently negative. In addition, respondents from middle-aged and older groups tended to rate their health less favorably than did younger respondents. These data support the hypothesis that general health ratings are valid measures of health status. General health ratings have been significantly (and often substantially) associated with measures of physical, mental, and social components of health status, supporting their validity as *general* measures of health. Finally, general health ratings were most often negatively associated with measures of use of medical care services, which also supported

Table 8

Summary of Published Studies of General Health Ratings and Age

General Health Measure[a]	Association[b] Sign	Statistic	Method	Sample Characteristics	Investigator(s)
General health rating	0	Not given	χ^2	Age deciles compared in general urban population, ages 21+	Schnore and Cowhig (1959-60)
General health rating	Not given	0.37	χ^2	Elderly volunteers in geriatric research program, ages 60-94	Heyman and Jeffers (1963)
Rating of present health	+	19.7[***]	χ^2	Volunteers in study of aging, ages 60-94	Maddox (1964)
Rating of current health	+	0.26	r	Elderly subjects in VA domiciliary unit, mean age 67.1	Desroches, Kaiman, and Ballard (1967)
Rating of past health	−	0.13	r	Elderly subjects in VA domiciliary unit, mean age 67.1	Desroches, Kaiman, and Ballard (1967)
Changes in health (past and present)	−	0.48[**]	r	Elderly subjects in VA domiciliary unit, mean age 67.1	Desroches, Kaiman, and Ballard (1967)
Poor health as reason not to work	−	Not given	%	Welfare mothers and working-class black women, age not reported	Cole and Lejeune (1972)
General health rating	−	79.97[**]	χ^2	General population; 3 age groups compared: 20-39, 40-64, 65+	Gaitz and Scott (1972)
Rating of present health	−	0.06[*]	r	Random sample of white health insurance plan enrollees, ages 46-71	Palmore and Luikart (1972)

[a] In all instances, data were obtained from the respondent.

[b] Abbreviations used in these columns:

[*] $p < 0.05$	% = proportions of population reported
[**] $p < 0.01$	χ^2 = chi-square value
[***] $p < 0.001$	r = product-moment correlation coefficient

their validity as measures of health status. Although general health ratings were associated with previous health and illness behavior in cross-sectional studies, establishing their validity as predictors of such behavior requires longitudinal data such as those available in the HIS.

DEMOGRAPHIC AND SOCIOECONOMIC CORRELATES OF GENERAL HEALTH RATINGS

Associations between general health ratings and education, income, occupational status, sex, and social class are not considered evidence of the validity of the ratings. Selected findings from 13 reports are summarized here for information purposes.

Inconsistent findings regarding relationships between general health ratings and demographic and socioeconomic variables may be due to true population differences in relationships or differences in survey instruments used (including their emphasis, reliability, and validity). In addition, correlations among demographic and socioeconomic variables have not been taken into account in studies of their relationships with general health ratings.

Education

Those reporting higher educational attainment tend to rate their health more favorably than others (Schnore and Cowhig, 1959-60; Osborn, 1973).

Income

Income and general health perceptions appear to be positively correlated. Hochstim, Athanasopoulos, and Larkins (1968) found that the higher the income (in a poverty area), the larger the proportion of respondents who rated their health positively. Although Fine (1975) reported a nonsignificant relationship between self-assessed health and current income, Schnore and Cowhig (1959-60) and Osborn (1973) reported significant positive associations between family income and reported health.

Occupational Level/Status

Heyman and Jeffers (1963) reported that a greater percentage of manual than nonmanual workers manifested negative changes in general health ratings. General health ratings tended to remain stable for nonmanual workers. Thompson and Streib (1958) compared changes in general health ratings for retirees and the employed. Both were equally likely to shift ratings; however, retirees were more likely to improve and the gainfully employed were more likely to decline in health perceptions.

Social Class

Of the three studies that addressed the relationship between general health ratings and social class, only Osborn (1973) reported a significant relationship;

people of higher socioeconomic status (SES) were more likely to report good health than those of lower SES. Similarly, Pratt (1971) reported trends showing that people of low SES with medium or low quality personal health practices had less favorable general health perceptions than those in the higher income groups with medium and low quality health practices. When quality of health practices was high, health ratings were equally favorable across SES groups. Garrity (1973) reported no relationship between general health ratings and SES.

Sex

Heyman and Jeffers (1963) and Schnore and Cowhig (1959-60) reported no significant relationships between general health ratings and sex. Maddox (1964), however, reported that optimism about health was greater for males; and Shanas, Townsend, Weddarburn, et al. (1968) found that men were more likely than women to rate their health favorably. The latter finding corresponds with National Health Interview Survey results (National Center for Health Statistics, 1976) indicating that men were more likely than women to report excellent health.

III. LITERATURE REVIEW: HEALTH PERCEPTIONS QUESTIONNAIRE

BACKGROUND

This section summarizes information regarding development and validation of Form II of the Health Perceptions Questionnaire (HPQ)[1] and results of previous surveys using the HPQ. This information served as a basis for deciding to add HPQ items to the HIS health status measurement battery and how to administer and score those items.

Two major questions are addressed: (1) How reliable and valid are HPQ scores? and (2) Are HPQ scores likely to contribute useful information about health over and above that already obtained with other HIS health status measures? Data relevant to these issues were gathered by Southern Illinois University School of Medicine for the National Center for Health Services Research.[2] Included were the results of cross-sectional surveys of approximately 2000 adults drawn from five general populations in Illinois and California between 1973 and 1975.[3] Table 9 summarizes the surveys and respondents. Demographic and socioeconomic characteristics of respondents varied considerably both within and between survey sites. Thus, results of these studies made it possible to address important measurement issues in groups of respondents differing widely in income and other variables of interest to the HIS.

CONTENT OF QUESTIONNAIRE AND ITEM SCORING

Form II of the HPQ was a standardized, self-administered survey instrument containing 36 favorably or unfavorably worded statements of opinion about personal health (see Table 10). HPQ items were drawn from questionnaire items used in previous research on general health and comments about health obtained during consumer surveys fielded by the developers of the HPQ. The items stressed general "health" and "feeling" states rather than particular components of health (i.e., physical, mental, social). Measurement emphasized six hypothesized constructs: perceptions of past, present, and future health (health outlook), resistance or susceptibility to illness, sickness orientation (tendency to accept illness as a part of life),

---

[1] Form II of the HPQ was developed following field tests and subsequent revisions of Form I. Only results of studies using Form II are reported here; Form I results are discussed by Ware and Karmos (1976).

[2] See Ware and Karmos (1976) for a report of project findings regarding general health perceptions and Ware (1976) for a summary of that report.

[3] One of the five field tests provided data only on long-term stability of HPQ scales; most analyses on reliability and validity were based on data from four field tests.

In all field tests, representative households were drawn using mixed sampling designs, and one or more adults from each household were selected as respondents. The rate of complete and usable interviews was approximately 37 percent for the field test that used mailout/mailback methods. Interviewer supervision of questionnaire administration resulted in rates of usable returns ranging from 62 to 83 percent. Complete and usable returns were received from approximately 46 percent in the field test that used interviewer dropoff and mailed return of questionnaires. Attempts to self-administer the HPQ were successful in all field tests. Missing responses to items in returned questionnaires were rare (a small fraction of 1 percent). Administration times averaged seven to nine minutes (slightly longer for disabled and disadvantaged respondents) in three field tests where times were recorded.

Table 9

SUMMARY OF RESPONDENT CHARACTERISTICS, FIVE FIELD TESTS
OF HEALTH PERCEPTIONS QUESTIONNAIRE

| Characteristics | Field Tests[a] | | | | |
	ESL	SAC	FPC	LAC	TRC
Sample Size	323	432	527	640	92
Sex (%)					
Male	19	22	36	63	16
Female	81	78	64	37	84
Race (%)					
White	10	97	(b)	65	90
Nonwhite	90	3	(b)	35	10
Age (years)					
Minimum	17	17	17	18	20
Maximum	88	84	84	92	80
Median	43	45	32	43	51
Family Income ($)					
Minimum	0	<2,000	0	0	0
Maximum	20,000+	20,000+	20,000+	30,000+	20,000
Median	5,400	11,900	12,000	9,500	9,000
Education (years)					
Minimum	3	3	6	0	5
Maximum	20+	20+	20+	20+	17
Median	11	12	14	12	11

[a] ESL = East St. Louis, Illinois.
SAC = Sangamon County, Illinois.
FPC = Family Practice Center, Springfield, Illinois.
LAC = Los Angeles County, California.
TRC = Tri-County, Illinois.
[b] Not available.

and health worry/concern. Variables other than health that might correlate with general health ratings were also measured, including tendency to reject the sick role and attitude toward going to the doctor.[4]

Each HPQ item was accompanied by five response categories: Definitely True, Mostly True, Don't Know, Mostly False, and Definitely False. Numbers ranging from one to five were assigned to these categories so that responses to most favorably and unfavorably worded HPQ health items defined a favorable health perception. For example, high scores for items in the Current Health scale indicated perceptions of good current health, and low scores indicated poor current health. Responses to other HPQ items were scored so that a high score defined a perception consistent with the name of the scale. High scores for items in the Resistance-Susceptibility to Illness scale, for example, indicated perceived resistance to illness, and low scores indicated perceived susceptibility. Table 11 presents item scoring information.

[4] Scaling and validity studies demonstrated that items emphasizing rejection of sick role and attitude toward going to the doctor reflected sick role propensity, as hypothesized, but shared little variance with other items. For these reasons, they were not considered measures of general health perceptions and construction of scales from these items is not discussed here.

Table 10

HEALTH PERCEPTIONS QUESTIONNAIRE ITEMS AND
HYPOTHESIZED GROUPINGS, FORM II

Hypothesized Item Grouping	Item Number[a]	Item
Current Health	1	According to the doctors I've seen, my health is now excellent.
	7[b]	The people I know seem to be healthier that I am.
	11	I feel better now than I ever have before.
	(15)	I am somewhat ill.
	(18)	I'm not as healthy now as I used to be.
	22	I'm as healthy as anybody I know.
	25	My health is excellent.
	(30)	I have been feeling bad lately.
	(34)	Doctors say that I am now in poor health.
	36	I feel about as good now as I ever have.
Prior Health	6[b]	I have been in bed a lot in the past because of illness.
	(19)	I was so sick once I thought I might die.
	26	I've never had an illness that lasted a long period of time.
	31	I have never been seriously ill.
Health Outlook	4[b]	I think my health will be better in the future than it is now.
	(9)	I will probably be sick a lot in the future.
	17	In the future, I expect to have better health than other people I know.
	23	I expect to have a very healthy life.

Table 10 (continued)

Hypothesized Item Grouping	Item Number[a]	Item
Health Outlook (continued)	27[b]	Most of the people I know will probably have fewer health problems than I will in the future.
	(35)	I think my health will be worse in the future than it is now.
Health Worry/ Concern	(10)	I never worry about my health.
	13	I worry about my health more than other people worry about their health.
	21	My health is a concern in my life.
	(28)	Others seem more concerned about their health than I am about mine.
Resistance- Susceptibility	(3)	I seem to get sick a little easier than other people.
	14	Most people get sick a little easier than I do.
	16	My body seems to resist illness very well.
	(32)	When there is something going around I usually catch it.
Sickness Orientation	12	Getting sick once in a while is a part of my life.
	29	I accept that sometimes I'm just going to be sick.
Rejection of Sick Role	2	I try to avoid letting illness interfere with my life.
	5	When I'm sick I try to keep it to myself.
	8	When I'm sick I try to just keep going as usual.
	33	When I think I am getting sick, I fight it.

Table 10 (continued)

Hypothesized Item Grouping	Item Number[a]	Item
Attitude Toward Going to the Doctor	(20)	I don't like to go to the doctor
	24	It doesn't bother me to go to a doctor.

[a]Items with numbers in parentheses scored so that Definitely True = 1, Mostly True = 2, Don't Know = 3, Mostly False = 4, and Definitely False = 5. Remaining items that satisfied scaling criteria scored so that Definitely True = 5, etc.

[b]These items did not satisfy scaling criteria and therefore were not retained to score scales.

Table 11

SCORING RULES FOR HPQ FORM II HEALTH SCALES

Scales	Items Summed for Scale Score[a]
Current Health	1 + 11 + 15 + 18 + 22 + 25 + 30 + 34 + 36
Prior Health	19 + 26 + 31
Health Outlook	9 + 17 + 23 + 35
Resistance-Susceptibility	3 + 14 + 16 + 32
Health Worry/Concern	10 + 13 + 21 + 28
Sickness Orientation	12 + 29

[a]Item numbers indicate placement in HPQ Form II; items should be scored as shown in Table 10 before computing scale scores.

SCALE CONSTRUCTION AND SCORING

Before the HPQ Form II items were scaled, they were grouped according to the specific perceptions they were hypothesized to measure (see Table 10). These groupings were based on results of factor analytic studies of HPQ Form I, reported elsewhere (Ware, Wright, and Snyder, 1974; Ware and Karmos, 1976).

The appropriateness of these item groupings (scales) was evaluated using expanded Likert (1932) and multitrait scaling criteria and factor analysis. Likert and multitrait criteria require that items correlate higher with the scale they are hypothesized to belong to than with other scales. Factor analysis aids in identifying item groupings that had not been hypothesized. (The logic and details of these methods are discussed in detail in Vol. I, Sec. II.)

Thirty-two HPQ Form II items satisfied the scaling criteria (six were items in the Rejection of Sick Role and Attitude Toward Going to the Doctor scales and are not discussed further here).[5] All hypothesized item groupings appeared as factors in the factor analyses in four field tests. Four items hypothesized to measure Current Health, Prior Health, and Health Outlook did not consistently meet the criteria across field tests and were eliminated from further analyses (see Table 10).

Interpretation of item content in each group that satisfied factor analytic criteria was the basis for assigning names to HPQ health scales: Current Health, Prior Health, Health Outlook, Health Worry/Concern, Resistance-Susceptibility to Illness, and Sickness Orientation. Five scales were balanced, containing both favorably and unfavorably worded items. The Sickness Orientation scale was constructed entirely from unfavorably worded items. Scores for this scale were more likely than those for balanced scales to be biased by tendencies to endorse or negate items regardless of their content (Ware, 1978). The extent of bias was evaluated during one field test, and results are discussed in a later section.

The score for each scale was computed for each respondent using the simple algebraic sum of scores for items that satisfied scaling criteria (see Table 11). As a result of the methods used to select items initially and revise them for Form II, it was not necessary to standardize or weight them for differences in variability or the extent to which they measured the scale construct. Items in each scale measured the construct defined by the scale more than they measured other constructs, and they measured that construct to about the same extent. Thus, item scores could simply be added to compute scale scores.

DISTRIBUTION OF SCALE SCORES

Table 12 summarizes descriptive statistics (means and standard deviations) and other information about scores computed from Form II scales in four field tests. Scale means were often close to the midpoint of the possible scale range, where the lowest possible score was the number of items in the scale and the highest possible score was five times that number. Standard deviations were usually one-fifth to one-sixth of the scale range. Thus, scale scores on the HPQ were symmetrically distributed, and there was considerable variability in scores. Means tended to be lower for HPQ health scales in the East St. Louis field test, in which a larger proportion of the sample was disadvantaged (and apparently less healthy).

---

[5] Factor matrices for items and matrices of item-scale correlations used in the multitrait scaling analysis are available elsewhere (Ware and Karmos, 1976).

Table 12

DESCRIPTIVE STATISTICS FOR HPQ FORM II HEALTH SCALES,
FOUR FIELD TESTS

Scales	No. of Items	Highest Possible Score	Midpoint of Scale Score	Means (Standard Deviations in Parentheses)			
				ESL[a]	SAC[a]	FPC[a]	LAC[a]
Current Health	9	45	27	27.6 (8.5)	32.9 (7.3)	32.7 (7.9)	32.7 (7.8)
Prior Health	3	15	9	9.4 (3.6)	10.5 (3.4)	10.5 (3.8)	10.3 (3.6)
Health Outlook	4	20	12	13.1 (2.4)	14.1 (2.6)	14.2 (2.8)	14.3 (2.8)
Resistance-Susceptibility	4	20	12	13.8 (3.0)	15.0 (2.6)	14.3 (3.2)	15.3 (2.6)
Health Worry/Concern	4	20	12	13.4 (2.8)	11.4 (2.7)	12.1 (2.8)	11.6 (3.1)
Sickness Orientation	2	10	6	7.1 (1.9)	6.7 (1.9)	7.0 (1.8)	6.2 (2.2)

[a]ESL = East St. Louis, Illinois.
SAC = Sangamon County, Illinois.
FPC = Family Practice Center, Springfield, Illinois.
LAC = Los Angeles County, California.

HOMOGENEITY OF SCALES

Homogeneity—average inter-item correlation—is a scale characteristic related to, but not the same as, the reliability of the scale score.[6] Estimates of homogeneity are useful because they indicate the extent to which scale items are reliable measures of the same construct. They are also unaffected by the number of items in a scale, so they serve as one basis for directly comparing scales that differ in length.

Homogeneity coefficients were computed for the six HPQ Form II health scales in each of four field tests; results are summarized in Table 13. These results indicate that items assigned to the same scale were homogeneous. Although most homogeneity coefficients were greater than 0.30 (a standard for evaluating homogeneity),[7] coefficients for Health Worry/Concern (a four-item scale) were nearly always below 0.30. With one or two exceptions, the lowest homogeneity coefficients were observed in the East St. Louis field test, which had the most disadvantaged population.

Table 13

SUMMARY OF HOMOGENEITY COEFFICIENTS FOR HPQ
FORM II HEALTH SCALES

	Coefficients		
Scales	High	Median	Low
Current Health	.57	.53	.48
Prior Health	.56	.47	.44
Health Outlook	.48	.43	.30
Resistance–Susceptibility	.50	.39	.25
Health Worry/Concern	.29	.26	.17
Sickness Orientation	.44	.39	.30

RELIABILITY

Internal-consistency reliability for HPQ scales was estimated independently in four field tests using Cronbach's (1951) Alpha coefficient and data obtained from all respondents. Test-retest reliability estimates were obtained for both item and scale scores in two field tests by computing product-moment correlations between scores obtained approximately six weeks apart from the same respondents (N = 183). Test-retest coefficients are biased downward whenever trait changes occur between administrations. (Methods of estimating reliability are discussed in greater detail in Vol. I, Sec. II.)

[6] Internal-consistency reliability is a function of scale length (number of items) and homogeneity (the average inter-item correlation). The relationship between internal-consistency reliability and the homogeneity of measurement is discussed further in Vol. I, Sec. II.

[7] Homogeneity coefficients of at least 0.30 are necessary to achieve an internal-consistency reliability coefficient of 0.90 with a 20-item scale.

Item Reliability

Table 14 summarizes test-retest reliability of scores computed from HPQ items. Items tended to be less reliable in the East St. Louis field test. The complete data reported by Ware and Karmos (1976) indicate that scores computed for 13 of the 26 HPQ health items were not sufficiently reliable for purposes of group comparisons in East St. Louis (according to the 0.50 standard suggested by Helmstadter, 1964); two of the 26 coefficients did not meet this standard in Sangamon County. This finding may reflect population differences in reliability or stability of the health constructs measured by these items.

Scale Reliability

Table 15 summarizes internal-consistency and test-retest reliability coefficients for HPQ scales across four field tests. (Coefficients for each field test are reported in App. A, Table A.1.) Almost without exception, reliability coefficients for scale

Table 14

SUMMARY OF TEST-RETEST RELIABILITY COEFFICIENTS
FOR SINGLE-ITEM HPQ SCORES,
TWO FIELD TESTS

Item Groupings/Field Tests[a]	Coefficients		
	High	Median	Low
Prior Health (3)			
East St. Louis	.59	.58	.50
Sangamon	.66	.60	.57
Current Health (9)			
East St. Louis	.62	.58	.38
Sangamon	.77	.64	.48
Resistance-Sysceptibility (4)			
East St. Louis	.64	.47	.38
Sangamon	.64	.56	.51
Health Outlook (4)			
East St. Louis	.45	.36	.28
Sangamon	.67	.63	.59
Sickness Orientation (2)			
East St. Louis	.48	(b)	.21
Sangamon	.62	(b)	.54
Health Worry/Concern (4)			
East St. Louis	.60	.48	.43
Sangamon	.57	.52	.46

[a]Number of single-item scores computed for each item grouping in parentheses.

[b]Not computed (only two coefficients).

Table 15

INTERNAL-CONSISTENCY AND TEST-RETEST RELIABILITY COEFFICIENTS,
HPQ FORM II HEALTH SCALES

Scales	No. of Items	Internal-Consistency Coefficients[a]			Test-Retest Coefficients[b]	
		High	Median	Low	ESL	SAC
Current Health	9	.92	.91	.89	.76	.86
Prior Health	3	.79	.73	.70	.67	.78
Health Outlook	4	.79	.75	.64	.54	.76
Resistance-Susceptibility	4	.80	.71	.58	.74	.73
Health Worry/Concern	4	.62	.60	.45	.60	.65
Sickness Orientation	2	.61	.59	.46	.42	.72

[a] Summary of coefficients across four field tests.

[b] ESL = East St. Louis, Illinois.
SAC = Sangamon County, Illinois.

scores exceeded 0.50 and were sufficiently reliable for group comparisons. In some instances, the Current Health scale was sufficiently reliable for individual comparisons (internal-consistency reliability coefficients equaled or exceeded 0.90, the generally accepted standard). Median coefficients (across field tests) ranged from a low of 0.59 for the two-item Sickness Orientation scale to a high of 0.91 for the nine-item Current Health scale.

Test-retest reliability coefficients for the six HPQ health scales ranged from a low of 0.42 to a high of 0.76 in East St. Louis and from 0.65 to 0.86 in Sangamon County. Differences between these coefficients and internal-consistency reliability estimates were equal to or greater than 0.10 for four of the six scales in one or both field tests where test-retest reliability was studied. For three of those scales (Resistance-Susceptibility, Health Worry/Concern, and Sickness Orientation), test-retest coefficients were greater than internal-consistency coefficients. For the Current Health scale, internal-consistency coefficients tended to be larger. Test-retest coefficients were higher than internal-consistency coefficients for three scales, suggesting that items in those scales contain unique reliable variance that tends to be stable over time. The opposite trend in differences between coefficients for the Current Health scale suggests that perceptions regarding current health status may be less stable over time. If this scale measures perceptions regarding current health, it should be sensitive to changes in scoring between administrations (four to six weeks). Such changes are plausible. The finding that test-retest coefficients tended to be lower than internal-consistency coefficients supports the validity hypothesis that the scale measures perceptions regarding health at the time of the interview.

Because reliability coefficients define the proportion of true score variance, they can be compared on a ratio scale. Such comparisons of test-retest coefficients for single-item and scale measures matched by construct clearly indicated substantial proportional increases in true score variance for scale measures. Reliability coefficients for HPQ scales represented an increase of approximately 40 to 90 percent in true score variance over the reliability of matched items (across field tests) in the least reliable and shortest scale (Sickness Orientation) and of approximately 30 to 40 percent for the most reliable and longest scale (Current Health).

Group Differences in Reliability of Scales

As noted earlier, HPQ scale scores tended to be least reliable in the East St. Louis field test. Because a large proportion of respondents in this field test were disadvantaged in terms of demographic and socioeconomic criteria, differences in reliability findings may have been due to personal characteristics of respondents. There were no obvious methodologic differences between field tests.

Groups differing in personal characteristics might appear to differ with respect to reliability of general health ratings for several reasons: (1) differences in homogeneity of scale items due to variations in item meaning across groups, (2) group differences in the stability of the underlying construct being measured, and (3) real differences in the reliability of ratings across population groups.

Factor analytic and multitrait scaling studies supported the appropriateness of HPQ item groupings across field tests. It was very unlikely that differences in reliability of HPQ scale scores could be explained entirely by differences in the meaning of items across population groups. Population differences in the stability of health perceptions may explain differences in test-retest reliability of HPQ scale scores but not those in their internal-consistency reliability. The third explanation—real differences in the reliability of ratings across population groups—seemed the most plausible.

To investigate this hypothesis further, internal-consistency reliability coefficients were computed independently for subgroups within each of the four field tests to estimate the proportion of true score variance defined by the six HPQ health scales (Ware and Karmos, 1976).[8]

Scale scores were slightly less reliable for respondents reporting less educational attainment, older age, and lower income; there were no differences in reliability between men and women. The largest differences observed were those between field tests. Actual group differences in reliability of general health ratings might have been detected had more disparate education and income groups been formed. Also, the most disadvantaged groups (e.g., respondents with less than an eighth-grade education) may yield the least reliable ratings. To address these issues, larger samples that permit study of larger, more disparate subsamples should be included in further research.

---

[8] Subgroups were defined as follows: (1) age (low = 40 years and younger, high = over 40); (2) education (low = 12 years or less, high = more than 12 years); (3) income (low = $7,000 or less, high = more than $7,000). Sample sizes for groups ranged from 50 to about 400 and exceeded 100 for most analyses.

STABILITY OF SCORES

The long-term stability of selected health perceptions measured by HPQ scales (Health Worry/Concern, Resistance-Susceptibility to Illness, Prior Health, and Current Health) was estimated from product-moment correlations between alternate forms of HPQ scales administered to the same respondents (N = 92) approximately two years apart. Intertemporal stability coefficients were positive and significant (p < 0.001, one-tailed test). These coefficients, which may be interpreted as the proportion of measured variance that remained stable, were 0.31 for Health Worry/Concern, 0.45 for Resistance-Susceptibility to Illness, 0.59 for Prior Health, and 0.62 for Current Health. Differences in the magnitude of these coefficients should not be interpreted as indicating differences in the stability of the traits over time because these scales were not equally reliable on a single administration. Also, these coefficients may have been underestimated slightly because of the noncomparability of alternate forms (Form I versus Form II of the HPQ) used to compute scores. Regardless, the substantial intertemporal correlation of general health perceptions suggests that repeated administrations (rather than one administration at the end of the HIS) of the HPQ will improve the precision of hypothesis testing in the HIS.

VALIDITY

Methods

In the absence of criteria against which to judge the validity of HPQ scales as measures of general health perceptions, two methods were used to obtain construct evidence of validity. The first involved studying associations among scales and the second, studying relationships that should exist if the scales measure the constructs they are supposed to measure. To ensure that the construct validity of the measures was being tested and that results would not be questionable on theoretical grounds, only those hypotheses based on well-accepted theory and well-documented relationships involving health status variables were tested.

To study associations among scales, matrices of correlations among the six health scales were factor analyzed to derive higher order factors, which were interpreted on the basis of the content of scale items with high loadings on the factors. Patterns of loadings across factors were then interpreted to evaluate the construct validity of each scale score—i.e., how well each scale measured each factor.[9] Because factor structures were similar across the four field tests, the proportion of reliable variance in each scale accounted for by each factor across field tests could be summarized without much loss of information.

The following hypotheses regarding factor analytic validity results were based on findings from Form I of the HPQ (see Ware, Wright, and Snyder, 1974; Ware and Karmos, 1976): (1) a major unrotated factor pertaining to perceptions of general health and another unrotated factor pertaining to health outlook would emerge; (2)

[9] For these validity studies, a matrix of product-moment correlations among the six health scales in the HPQ was computed for each of four field tests. Factors were extracted and rotated to orthogonal simple structure (see Vol. I, Sec. II for further description of methods).

HPQ scales primarily measuring past and present general health (Prior Health and Current Health) would correlate highly (± 0.40 or greater) with the same rotated health factor and with the factor pertaining to health outlook; (3) the remaining scales (Health Outlook, Health Worry/Concern, Resistance-Susceptibility to Illness, and Sickness Orientation) would correlate highly with the rotated factor defining health outlook and not with the rotated factors pertaining to present and prior health; (4) Resistance-Susceptibility to Illness and, to a lesser extent, Health Worry/Concern would have noteworthy secondary loadings on factors pertaining to current and prior health. (These hypotheses were based on studies of interrelationships among HPQ Form I scales and are consistent with the notion that perceptions of resistance-susceptibility and health worry/concern are intervening variables between perceptions of prior and current health and those of future health.)

The second method used to examine construct validity involved evaluating relationships between the HPQ scale scores and 11 other measures of health, seven measures of health and illness behavior, and age. Table 16 defines the validity variables studied and presents specific hypotheses regarding their associations with favorable health perceptions. Briefly, the following hypotheses were studied: that favorable health perceptions would be positively associated with positive definitions of health (e.g., psychological well-being), negatively associated with negative definitions of health (e.g., bed days), negatively associated with illness behavior (e.g., number of doctor visits), and would not be associated with health behavior (e.g., checkups);[10] that perceptions of prior, current, and future health and resistance-susceptibility would be negatively correlated with age; and that perceptions of sickness orientation and health worry/concern would increase with age.

Higher Order Factor Structure

Factor analytic studies of the construct validity of HPQ scales reported by Ware and Karmos (1976) were based on the principal factor method and a matrix of correlations among the six HPQ health scales and the two HPQ scales measuring patient role propensity (Rejection of Sick Role and Attitude Toward Going to the Doctor). Three higher order [11] factors were studied that collectively accounted for approximately 70 percent of the total reliable variance in eight scales across four field tests. Results indicated that the six HPQ health scales overlapped very little with the two patient role propensity scales and that the six scales differed considerably in factor content. Thus, a different interpretation for each scale score appeared necessary. However, construct validity results indicated that the interpretation of each health scale, based on relationships with the higher order factors, was substantially the same across all four field tests.

Correlations among the six HPQ health scales were reanalyzed to (1) determine the extent of overlap with a general health perceptions factor, (2) determine the amount of unique information about health status contributed by each scale, and (3) increase understanding of the meaning of each scale by studying its construct validity in relation to a larger number of factors.

---

[10] The hypothesis regarding no association between favorable health perceptions and health behavior was based on previous findings (Tissue, 1972; Ware, Wright, and Snyder, 1974).

[11] The term "higher order" refers to factor analyses in which the variables analyzed were factors (scales) that resulted from a previous factor analysis; in the present case, the factors used were derived from factor analysis of questionnaire items.

Table 16

DEFINITION OF VALIDITY VARIABLES AND DIRECTION OF HYPOTHESIZED
RELATION WITH FAVORABLE HEALTH PERCEPTIONS

Category/Variable	Definition	Hypothesized Relationship with Favorable Health Perceptions
Health Status		
General health item	Rating of general health as excellent, good, fair, or poor	+
Role activity limitations	Number of days during the previous year that the respondent was too sick to perform usual activities	−
Any bed days	Whether respondent spent all or part of any day in bed because of illness or injury during the previous two months	−
Total bed days	For those respondents with one or more bed days during the previous two months, the number of days in which all or part of the day was spent in bed because of illness or injury	−
Sickness	Presence or absence of sickness during the previous two months	−
Chronic problems	Presence or absence of any continuous or recurring health problem (one that is present all the time or that makes the respondent ill from time to time)	−
Pain	Four-choice rating of how often pain was experienced during the previous year	−
Feeling poorly	Number of days during the previous two months on which the respondent did not feel as well as usual	−
Health worry	Four-choice rating of worry about health during the previous year	−
Psychological well-being	10-item scale to measure positive versus negative psychological well-being[a]	+
Reason for last doctor visit	Whether the last doctor visit was for sickness/injury	−

Table 16 (continued)

Category/Variable	Definition	Hypothesized Relationship with Favorable Health Perceptions
Health and Illness Behavior		
Number of doctor visits	Number of doctor visits during the previous year	−
Dental visit	Whether the respondent saw a dentist during the previous year	−
Check-up	Whether a doctor was seen for a check-up or physical examination, when not sick, during the previous year	0
Doctor visit when needed	For those respondents reporting sickness, injury, or accident during the previous two months, whether a doctor was seen	−
Recency of care	Number of months since last doctor visit for any reason	+
Hospitalization	Whether respondent stayed over-night in a hospital (other than for childbirth) during the pre-vious year	−
Compliance	Whether the respondent had been following doctor's orders exactly, for those on medical regimens	−

[a]From Bradburn, 1969.

Results from the additional analyses regarding overlap between scales and a general health factor are presented in Table 17, which summarizes the highest (H) and lowest (L) percentages of reliable variance in each scale accounted for by one to six factors. Two trends in Table 17 are most important. First, the last row indicates that the first unrotated factor accounted for a very large percentage of the total reliable variance defined by the six scales (59 to 62 percent across four field tests). In terms of variance accounted for, this large general factor was most related to the Current Health scale (72 to 73 percent of the variance in Current Health overlapped with the first unrotated factor across the four field tests). It was also substantially related to the Prior Health, Resistance-Susceptibility, and Health Outlook scales (one-third to over one-half of the variance in these scales was accounted for by the general factor across field tests). Health Worry/Concern and Sickness Orientation scales overlapped much less with the first unrotated factor.

Second, the additional percentage of variance in the six scales explained by the extraction of more factors diminished rapidly; the first four factors explained all but 6 to 10 percent of the total reliable variance. However, six factors were necessary to account for all of the reliable variance in the six scales. More than 5 percent of the reliable variance in two scales (Current Health and Resistance-Susceptibility) was unaccounted for by the five-factor solution. Whether this variance is trivial or important in defining health status could not be determined from these data. The results did indicate that each HPQ health scale contains some reliable information

Table 17

SUMMARY OF PERCENTAGES OF MEASURED VARIANCE IN SIX HPQ SCALES
EXPLAINED BY ONE TO SIX FACTORS, FOUR FIELD TESTS

	Number of Factors[a]											
	1		2		3		4		5		6[b]	
Scales	H	L	H	L	H	L	H	L	H	L	H	L
Current Health	73	72	80	75	84	76	86	82	89	84	91	89
Prior Health	40	34	66	46	73	56	74	67	79	69	79	70
Health Outlook	55	36	57	55	69	57	70	60	74	64	79	64
Resistance- Susceptibility	55	46	62	50	62	51	62	52	79	53	80	58
Health Worry/Concern	29	13	36	16	54	33	58	44	62	44	62	45
Sickness Orientation	20	11	45	18	57	19	58	45	60	46	61	46
Proportion of total[a]	62	59	75	72	84	82	94	90	98	96	100	100

[a]Highest (H) and lowest (L) percentage observed across four field tests.

[b]Percentages accounted for by six factors were equal (within rounding error) to the internal-consistency reliability estimates, which were used as communality estimates.

[c]Total = total reliable variance.

not shared with the other scales. Thus, there is reason to score and interpret the scales separately in HIS analyses.

The four-factor solution was evaluated to increase understanding of the meaning of the six scales in relation to a larger number of higher order factors.[12] The four-factor solution provided the smallest number of factors that could be studied without ignoring more than 10 percent of the reliable variance and was the most consistent solution across the four field test sites. The four-factor solution was also the largest number of factors that could be studied and still retain a clear picture of major sources of common factor variance (factors that accounted for a noteworthy amount of variance in two or more scales). The rotated four-factor solutions for each field test are presented in Tables A.2–A.5 in App. A. Table 18 summarizes the percentages of reliable variance in the six HPQ health scales accounted and not accounted for by the four rotated higher order factors. With one exception, results were similar across field tests and could be summarized in one table; the exception is discussed below.

As expected, rotation of factors in addition to those previously described by Ware and Karmos (1976) substantially reduced the amount of unexplained variance in several of the HPQ health scales (Prior Health, Health Worry/Concern, and Sickness Orientation). Before this improvement is discussed, however, note that information about validity was lost because the two patient role propensity scales were not included. In the factor analysis of eight scales reported previously (Ware and Karmos, 1976), both the Resistance-Susceptibility and Health Worry/Concern scales overlapped considerably (6 percent or more of the reliable variance) with a patient role propensity factor not included in the reanalysis. This variance should be kept in mind when scores on these two scales are interpreted.

The four rotated factors were interpreted from their patterns of correlations with the six HPQ health scales (see Tables A.2–A.5 in App. A). Assuming, from the content of scale items, that each scale measures a perceptual construct similar to the scale name, interpretation of the four factors was straightforward. With few exceptions, each factor had one very high correlation with one of the six scales and low correlations with the others. Nearly all exceptions to this pattern of results across field tests were accounted for by the second factor, which was associated with very high loadings for both the Health Outlook and Current Health scales in all field tests. Thus, these scales share much more common factor variance than the other scales in the matrix. Apparently, respondents viewed their chances of good future health in a manner consistent with their evaluation of present health status. From 87 to 89 percent of the variance in three of the other scales (Prior Health, Health Worry/Concern, Sickness Orientation) was accounted for by one factor.

The most complicated scale was the Resistance-Susceptibility scale, which did not have the simple pattern of loadings observed for the other five scales. As shown in the third row of Table 18, from 8 to 27 percent of the variance in this scale was accounted for by each of the four rotated factors. Thus, a regression of the Resistance-Susceptibility scale on the four factors would reveal highly significant regression coefficients for all factors in all field tests. This presents problems in interpreting Resistance-Susceptibility scale scores; if differences between groups

---

[12] As previously noted, six factors were necessary to account for all of the reliable variance in the six scales. Four factors accounted for 90 to 94 percent of the reliable variance across field tests (see col. 4 in Table 17).

Table 18

PERCENTAGES OF RELIABLE VARIANCE IN SIX HPQ SCALES ACCOUNTED
AND NOT ACCOUNTED FOR BY FOUR DERIVED FACTORS

	Reliable Variance	Current Health/ Health Outlook	Prior Health	Health Worry/Concern	Sickness Orientation	Unknown
				Sources of Reliable Variance		
Current Health	91	65	19	5	0	11
Prior Health	74	5	88	1	1	5
Resistance-Susceptibility	70	26	3	27	11	28
Health Outlook	73	79	1	3	5	12
Health Worry/Concern	55	4	2	87	2	5
Sickness Orientation	55	4	2	2	89	3

were observed in scores, they could be due to differences on any one or more of the higher order factors. The percentage of reliable variance not explained by the four factors was also largest (28 percent, on the average) for the Resistance-Susceptibility scale. Ware and Karmos (1976) reported that a small but significant amount of this variance was accounted for by differences in patient role propensity, leaving about one-fourth of the reliable variance unaccounted for.

As noted earlier, the higher order factor solutions were similar across field tests. The major exception was a much higher degree of overlap between the Prior Health factor and the Current Health scale in the study of urban black adults in East St. Louis.[13] This finding is consistent with the hypothesis that the meaning of general health ratings may vary significantly across groups differing in demographic and socioeconomic characteristics.

This discussion has emphasized the extent of overlap between HPQ scales and the higher order factors. The direction of the associations involved is also an important consideration in interpreting the scales. (Signs associated with correlations between the scales and higher order factors are presented in Tables A.2–A.5, App. A.) Without exception, signs associated with high factor loadings for scales were the same across field tests. Associations among the Prior Health, Current Health, Health Outlook, and Resistance-Susceptibility scales and factors defined by these scales were all positive. Associations between all of these scales and the Health Worry/Concern and Sickness Orientation scales and the factors defined by the latter two scales were all negative. The correlation between the latter two scales was positive in all field tests.

HPQ Scales and Health-Related Variables

Results regarding associations between HPQ scales and 21 health-related variables are summarized in Tables 19 and 20. Included among these variables were 11 measures of health status, seven measures of health and illness behavior, two measures of patient role propensity, and age. Detailed results for each of the four field tests are presented in App. A (see Tables A.6–A.11).

For both Current Health and Prior Health scales, relationships were significant and positive for variables defining favorable health states (e.g., psychological well-being) and negative for variables defining poor health (e.g., bed days), as hypothesized. Many of these coefficients were moderately high (above 0.40), indicating substantial relationships. For Current Health, correlations were higher for measures of feelings or perceptions (e.g., general health item, pain, and worry) than for more directly observable phenomena (e.g., bed days and role activity limitations).

With few exceptions, the other HPQ scales correlated with other health variables across field tests as hypothesized. Correlations were significant and higher in absolute magnitude more often for Health Outlook and Resistance-Susceptibility to Illness than for Health Worry/Concern and Sickness Orientation. Significant coefficients for the latter two scales and other health variables defined negligible

---

[13] Correlations between the Current Health scale and the Prior Health factor (see col. 2 in Tables A.2–A.5, App. A) ranged from 0.31 to 0.42 in the Sangamon County, Family Practice, and Los Angeles County field tests. This correlation was 0.60 in East St. Louis. When squared and adjusted for differences in reliability across field tests, these results indicated that from 10 to 18 percent of the reliable variance in the Current Health scale was accounted for by the Prior Health factor in three field tests; in East St. Louis this figure was 41 percent.

Table 19

SUMMARY OF VALIDITY COEFFICIENTS FOR PRIOR HEALTH, CURRENT HEALTH, AND
HEALTH OUTLOOK SCALES IN RELATION TO HEALTH STATUS, HEALTH AND
ILLNESS BEHAVIOR, PATIENT ROLE PROPENSITY, AND AGE

Validity Variables	Prior Health		Current Health		Health Outlook	
	No.[a]	r[b]	No.[a]	r[b]	No.[a]	r[b]
Health Status						
General health item	1/1	.50[c]	1/1	.80[c]	1/1	.58[c]
Role activity limitations	4/4	-.36[c]	4/4	-.46[c]	4/4	-.24[c]
Any bed days	1/1	-.21[c]	1/1	-.37[c]	1/1	-.23[c]
Total bed days	1/1	-.30[c]	1/1	-.38[c]	0/1	-.16
Sickness	1/1	-.18[c]	1/1	-.24[c]	0/1	-.13[c]
Chronic problems	1/1	-.32[c]	1/1	-.43[c]	1/1	-.34[c]
Pain	1/1	-.39[c]	1/1	-.55[c]	1/1	-.35[c]
Feeling poorly	1/1	-.19[c]	1/1	-.28[c]	1/1	-.13[c]
Health worry	1/1	-.44[c]	1/1	-.58[c]	1/1	-.33[c]
Psychological well-being	1/1	.19[c]	1/1	.39[c]	1/1	.28[c]
Reason for doctor visit	1/1	.11[c]	1/1	.21[c]	0/1	.09
Health and Illness Behavior						
Number of doctor visits	3/3	-.32[c]	3/3	-.42[c]	3/3	-.16[c]
Dental visits	1/3	-.02	1/3	.03	0/3	.06
Check-up	0/3	-.03	1/3	-.05	1/3	-.01
Doctor visit when needed	1/1	-.27[c]	1/1	-.47[c]	1/1	-.26[c]
Recency of care	2/2	.15[c]	1/2	.16[c]	0/2	.06
Hospitalization	3/3	-.28[c]	3/3	-.30[c]	2/3	-.11[c]
Compliance	0/1	-.07	1/1	-.14[c]	0/1	-.06
Patient Role Propensity						
Rejection of Sick Role	0/4	.02	0/4	-.02	0/4	.04
Attitude Toward Going to the Doctor	1/4	.01	2/4	.09	1/4	.07
Demographic Variable						
Age	4/4	-.19[c]	4/4	-.26[c]	4/4	-.23[c]

[a] Number of significant coefficients in the hypothesized direction in relation to number of field tests.

[b] Product-moment correlation not corrected for attenuation. In those instances where more than one correlation was independently computed, the table entry is the median coefficient across field tests.

[c] $p < 0.01$ (one-tailed test).

Table 20

SUMMARY OF VALIDITY COEFFICIENTS FOR RESISTANCE-SUSCEPTIBILITY,
HEALTH WORRY/CONCERN, AND SICKNESS ORIENTATION SCALES IN
RELATION TO HEALTH STATUS, HEALTH AND ILLNESS BEHAVIOR,
PATIENT ROLE PROPENSITY, AND AGE

Validity Variables	Resistance-Susceptibility		Health Worry/Concern		Sickness Orientation	
	No.[a]	r[b]	No.[a]	r[b]	No.[a]	r[b]
Health Status						
General health item	1/1	.50[c]	1/1	-.38[c]	1/1	-.16[c]
Role activity limitations	4/4	-.33[c]	4/4	.19[c]	3/4	.14[c]
Any bed days	1/1	-.24[c]	1/1	.10[c]	1/1	.14[c]
Total bed days	0/1	.00	0/1	.07	0/1	.01
Sickness	1/1	-.21[c]	1/1	.12[c]	1/1	.12[c]
Chronic problems	1/1	-.26[c]	1/1	.10[c]	1/1	.14[c]
Pain	1/1	-.41[c]	1/1	.24[c]	1/1	.18[c]
Feeling poorly	1/1	-.21[c]	1/1	.14[c]	0/1	.09
Health worry	1/1	-.42[c]	1/1	.42[c]	1/1	.20[c]
Psychological well-being	1/1	.18[c]	1/1	-.12[c]	0/1	-.11[c]
Reason for doctor visit	1/1	.17[c]	0/1	-.05	0/1	.01
Health and Illness Behavior						
Number of doctor visits	3/3	-.32[c]	3/3	.20[c]	1/3	.10
Dental visits	0/3	-.02	0/3	-.01	0/3	.00
Check-up	0/3	-.01	0/3	-.04	0/3	.02
Doctor visit when needed	1/1	-.21[c]	1/1	.17[c]	0/1	.09
Recency of care	2/2	.15[c]	2/2	-.15[c]	0/2	.00
Hospitalization	3/3	-.26[c]	1/3	-.07	0/3	.02
Compliance	0/1	-.06	1/1	.16[c]	0/1	-.10
Patient Role Propensity						
Rejection of Sick Role	3/4	.16[c]	4/4	-.21[c]	1/4	-.10
Attitude Toward Going to the Doctor	1/4	.03	1/4	.03	0/4	.00
Demographic Variable						
Age	1/4	.09	3/4	-.10[c]	2/4	-.06

[a] Number of significant coefficients in the hypothesized direction in relation to number of field tests.

[b] Product-moment correlation not corrected for attenuation. In those instances where more than one correlation was independently computed, the table entry is the median coefficient across field tests.

[c] $p < 0.01$ (one-tailed test).

relationships in many instances. Current Health and Prior Health scales both correlated about the same with measures of whether any bed days occurred and the number of bed days (for those who reported one or more). This was not the case for the other four HPQ scales, which related more to whether respondents reported any bed days than to the number. Thus, these scales may be more sensitive to whether disability occurred than to amount.

For all six HPQ health scales, correlations with measures of health and illness behavior were lower than correlations with health status variables. The hypothesized significant and negative relationships between both Current Health and Prior Health and illness behavior were observed; they were higher for Current Health. Some of the latter negative coefficients were moderate, indicating substantial associations between general health ratings and use of health care services. Current Health (but not Prior Health) was also significantly related to whether respondents reported compliance. Resistance-Susceptibility, Health Outlook, and Health Worry/Concern were most often correlated with illness behavior as hypothesized; Sickness Orientation was not. Health Worry/Concern was significantly correlated with compliance (those who worried were more likely to comply).

Current Health, Prior Health, and Health Outlook scale scores tended to decrease significantly with age in all field tests, as hypothesized. Scores for the other three scales were related to age in the hypothesized direction, although coefficients were not consistently significant across field tests. When examined for curvilinearity, relationships between age and HPQ scale scores tended to be linear when significant. Weak curvilinear relationships occurred because very favorable health perceptions were rarely observed for those over 70, whereas the full range of scores was observed for younger respondents.

Three scales (Prior Health, Current Health, and Health Outlook) were not significantly related to the measure of sick role rejection in any of the four field tests. Those scoring high on the Health Worry/Concern scale were less likely to reject the sick role and those scoring high on the Resistance-Susceptibility scale tended to reject the sick role (in three of four field tests). Thus, those who are worried about their health or perceive themselves susceptible to illness have greater propensity to assume the patient role. Associations between the six HPQ health scales and the other patient role propensity measure (attitude toward going to the doctor) were not consistently significant across field tests. Attitude toward going to the doctor tended to be more favorable for those scoring high on Current Health in two field tests; none of the other scales correlated significantly with this measure in more than one field test. Differences in patient role propensity, therefore, do not warrant consideration in interpreting scores for four of the HPQ health scales (Prior Health, Current Health, Health Outlook, and Sickness Orientation). Patient role propensity (measured in terms of sick role rejection) should be considered when interpreting Health Worry/Concern and Resistance-Susceptibility scales.

The analyses of HPQ scales and health and illness behavior described above and those reported by Ware and Karmos (1976) did not take into account correlations among the scales. Additional analyses of data from two field tests (Sangamon County and East St. Louis) were performed (for this review) to test whether more than one HPQ scale would be useful in predicting number of doctor visits during the year before the interview. All zero-order correlations between the six HPQ

health scales and number of doctor visits were statistically significant in both field tests (see Tables A.6–A.11). When correlations among the scales were taken into account in a regression analysis, however, only Current Health and Prior Health were associated with significant ($p < 0.05$) regression coefficients, both negative. In East St. Louis, the Current Health and Resistance-Susceptibility scales were associated with significant coefficients in the regression analysis, also negative. Because previous behavior probably affected ratings, these studies allowed only weak tests of predictive validity.

SOCIODEMOGRAPHIC CORRELATES OF SCALES

Associations among the six HPQ health scales and four demographic and socio-economic variables are summarized in Table 21 for four field tests. These associations are not considered evidence of the validity of the measures; they are presented for information purposes.

Current Health, Prior Health, and Health Outlook scale scores tended to increase with increases in education, income, and social class. Health Worry/Concern scores were consistently lower for those reporting higher income but not for those higher in educational attainment or social class. Relationships between other HPQ scale scores and education, income, and social class were not consistent across field tests. Results regarding differences in general health ratings for men and women were not consistent across field tests. Although not replicated, trends indicated that men rated their previous health and ability to resist illness more favorably than women and that women held a more favorable outlook regarding future health.

Table 21

SUMMARY OF SIGNIFICANT CORRELATIONS BETWEEN HPQ FORM II HEALTH
SCALES AND DEMOGRAPHIC AND SOCIOECONOMIC VARIABLES

								Variables						
	Education			Income			Social Class[b]			Sex[c]				
Scales	No.[a]	L[a]	H[a]	No.	L	H	No.	L	H	No.	L	H		
Current Health	4	$.20^d$	$.32^d$	4	$.20^d$	$.27^d$	2	$-.22^d$	$-.24^d$	0	NS	NS		
Prior Health	3	NS	$.22^e$	4	$.14^d$	$.21^d$	2	$-.16^f$	$-.16^d$	1	NS	$.16^f$		
Health Outlook	4	$.14^f$	$.28^f$	3	NS	$.14^d$	2	$-.12^f$	$-.16^f$	1	NS	$-.11^e$		
Resistance-Susceptibility	2	NS	$.29^d$	3	NS	$.24^d$	2	$-.11^e$	$-.16^f$	1	NS	$.12^f$		
Health Worry/Concern	1	NS	$-.11^f$	4	$-.10^e$	$-.14^d$	1	NS	$.24^d$	0	NS	NS		
Sickness Orientation	1	NS	$-.08^e$	1	NS	$-.16^f$	0	NS	NS	0	NS	NS		

[a]No. = number of significant coefficients across four field tests, L = lowest, and H = highest coefficients (absolute magnitude).

[b]Correlations with social class were computed in only two field tests (Sangamon County and East St. Louis). Social class was scored according to the Hollingshead Index with five ordered categories ranging from I = High to V = Low.

[c]Where sex scored 1 = female, 2 = male; a positive correlation indicates men scored higher than women on that scale.

[d]$p < 0.001$, two-tailed test.

[e]$p < 0.05$, two-tailed test.

[f]$p < 0.01$, two-tailed test.

IV. SUMMARY AND DISCUSSION

Taken at face value, measures of general health perceptions differ from other health status measures in that they do not specify one or more components of health (physical, mental, or social). Rather, respondents are asked only for an assessment of their "health." In theory, this difference in measurement strategy makes it possible to tap both the objective information people have about their health status and their evaluation of that information.

The issue at hand is whether measures of these perceptions yield reliable and valid information about health status. The extent to which general health ratings contribute information about health status beyond that available from HIS measures already being fielded (e.g., functional status measures of physical health as well as measures of mental health, social health, and disability days) is also important to HIS analyses.

These issues were addressed by reviewing published literature relevant to the reliability and validity of general ratings as well as recently completed studies of the Health Perceptions Questionnaire (HPQ) constructed by Ware and Karmos (1976). In certain instances, additional analyses of their data were also performed. Specifically, the following questions were addressed:

1. How have measures of general health perceptions been used?
2. How have general health perceptions been defined and measured?
3. Are general health perceptions reliable and stable over time?
4. For what kinds of conclusions and predictions about health status are general health perceptions valid?

TRENDS IN THE LITERATURE BEFORE THE HPQ

Forty publications pertaining to general health perceptions published during the past 25 years were identified. General health ratings are among the more frequently used measures of health status. Even so, their reliability and validity have not been rigorously evaluated. In addition, studies have differed widely in scoring techniques and methods of statistical analysis. Thus, comparison of results across investigations was difficult. With few exceptions, general health perceptions have been scored from single-item measures for which reliability estimates were not published.

Two publications included reliability estimates for general health rating scales; three reported estimates of stability over time (two to three years). These results suggest that ratings may be reliable enough for group comparisons and stable enough over time to warrant their use in a repeated-measures design like that adopted in the HIS. However, published results do not warrant generalization to the HIS.

Published research rarely addressed the validity of general health rating scores—i.e., what they measure and how they should be interpreted. Published associations between ratings and other measures of health status and health and illness behavior, however, proved useful in assessing the validity of ratings. In most

instances, the observed pattern of associations was what would be expected if general health ratings are a measure of *general* health, supporting their construct validity. General health ratings were significantly associated with the following variables across studies: functional limitations; physical health (e.g., bed days, number of illnesses or problems present); subsequent mortality; mental health (e.g., psychological well-being, depression); social health (e.g., social participation); physicians' assessments of respondents' health; use of health care services; ratings of the health of others; health knowledge; and personal health practices.

Many statistical tests used in published studies, particularly those that treated general health ratings as nominal data, probably were not as powerful as they might have been. Reported studies may have missed or underestimated associations between general health ratings and other health variables, or overlooked differences in general health among diverse population groups. Nevertheless, there is evidence that general health ratings do reflect health. Without reliability data, however, published validity findings cannot be evaluated to determine the strength of the associations involved.

STUDIES OF THE HPQ

Findings from studies of the HPQ published by Ware and Karmos (1976) and additional analyses of their data made it possible to address certain measurement issues more directly. These studies were fielded specifically to provide data necessary for construction of general health rating scales and to study reliability and validity of those scales in four general populations.

Construction of HPQ Scales

Studies of an early version of the HPQ (Ware, Wright, and Snyder, 1974) indicated that respondents are able to distinguish among perceptions of what their general health has been, what it is now, and what they think it will be in the future. On the strength of this finding and results published by others, a taxonomy of general health perceptions was defined and items were constructed to measure the following constructs: Prior Health, Current Health, Health Outlook, Resistance-Susceptibility to Illness, Health Worry/Concern, and Sickness Orientation.[1]

Rigorous scaling criteria were used to verify item groupings hypothesized to measure each of these constructs, and attempts to construct scales from HPQ items were successful. Retained items in hypothesized groupings were clearly shown to measure primarily the same constructs. Item groupings as hypothesized were empirically validated across independent field tests, providing strong support for the taxonomy of health perceptions on which questionnaire construction was based. The consistency of findings across populations differing widely in age, level of educational attainment, income, and race supported the generalizability of conclusions regarding the scales.

Scale scores had sufficient variance to warrant their use in general populations and were roughly symmetrically distributed, facilitating their use in conjunction

---

[1] Construction of items and scales to measure two nonhealth constructs (Rejection of the Sick Role and Attitude Toward Going to the Doctor) were not discussed in the literature review in this volume.

with the more powerful statistical methods. Achievement of these and other goals (discussed below) were essential to use of the HPQ as a *general* health measure.

Reliability and Stability

Studies of reliability based on test-retest and internal-consistency methods indicated that HPQ scores were sufficiently reliable for purposes of group comparisons and that single-item scores often were not. Because the majority of studies of general health ratings published in the last 25 years were based on single-item measures, the strength of most associations between general health perceptions and other variables, as reported in the literature, has probably been underestimated.

Internal-consistency and test-retest reliability coefficients for HPQ scales tended to be lower in disadvantaged populations, although the amount of true score variance defined by scales was sufficient to warrant their use in all populations studied. Reliability coefficients were lower for disadvantaged respondents. Furthermore, although test-retest coefficients were lower than internal-consistency coefficients in all field tests, they were considerably lower for disadvantaged populations. These trends suggest that general health perceptions are less reliable at any one time and less stable over time for the disadvantaged.

The stability of four HPQ constructs (Current Health, Prior Health, Resistance-Susceptibility to Illness, and Health Worry/Concern) was evaluated over a two-year period in one field test site, and stability was sufficient to warrant their use in repeated-measures research designs. Use of general health ratings in a repeated-measures design would improve precision of estimates, or would permit a reduction in the sample size necessary to detect group differences. The finding that general health perceptions tended to be stable over time was consistent with results published for other measures (Suchman, Phillips, and Streib, 1958; Heyman and Jeffers, 1963; Maddox, 1964; Reynolds, Rushing, and Miles, 1974); HPQ findings suggest that the long-term stability of these perceptions has been underestimated. Lower estimates in published studies probably resulted from the lower reliability (at each time) of the single-item scores or from the dichotomous scoring methods used.

Favorable results regarding the reliability and stability of general health ratings are particularly important given the prejudice against use of such ratings on these grounds.

Validity: Factor Analysis of HPQ Scales

Factor analytic studies of the construct validity of the HPQ reported by Ware and Karmos (1976) indicated that the health scales tended to cluster much as hypothesized and that they overlapped very little with patient role propensity (a nonhealth-related factor hypothesized to influence conscious decisionmaking regarding illness behavior). Additional analyses of their data indicated that each scale contributed some unique information about health perceptions and that scores on five scales (all but Resistance-Susceptibility) could be interpreted straightforwardly.

These results constituted strong support for the construct validity of HPQ scale scores: Each scale tends to measure the perceptual construct it was intended to measure to a greater extent than it measures constructs it was not intended to

measure. The findings support construct validity because factor structure was fairly simple to interpret and results were nearly identical across independent field tests. One exception was the factor structure of the Current Health scale in the East St. Louis field test (a population of urban blacks). In that field test, the Current Health scale overlapped substantially more with the factor defining perceptions regarding prior health. This trend suggests that the validity of scales to measure general health perceptions may differ substantially across groups differing in demographic and socioeconomic characteristics.

The results were also an important step in the difficult process of determining the meaning that should be assigned to scores computed from these measures and, therefore, the appropriate uses of the scales. Factor analytic evidence of construct validity also called attention to the fact that some HPQ scales, as do most health status measures, assess more than one health construct. Although five of the six scales consistently had a major source of reliable variance, none had a single source. For example, a high Current Health score indicated favorable perceptions of health in the present, a favorable health outlook, and, to a lesser but noteworthy extent, an absence of worry regarding health. The scale with the most complex factor content was Resistance-Susceptibility, which overlapped significantly with all four higher order factors (Prior Health, Current Health/Health Outlook, Health Worry/Concern, and Sickness Orientation).[2]

Findings also suggest that a global health perceptions index constructed from the six HPQ scales would be very reliable and would capture the majority of reliable variance in all field tests. A noteworthy amount of reliable variance would be lost, however; about 40 percent with one global index, about 25 percent with two, and so on (see Table 17). Further research is necessary to determine the practical implications of these tradeoffs in terms of defining health status and predicting health and illness behavior.

Validity: HPQ Scales and Other Health-Related Variables

Relationships among the scales and between the scales and higher order factors pertaining to health status supported the hypothesis that the scales measured health status. If the scales are valid measures of general health, they should also be substantially related to measures of physical and mental health and to behavioral variables (e.g., use of health care services) that are, in theory, influenced by health status.

Other Health Status Variables. Relationships between the six HPQ scales and the 11 other health status variables studied were strong enough to conclude that the scales measure health status and weak enough to indicate that they may contribute unique information about health. As hypothesized, perceptions tended to be more unfavorable in conjunction with increases in role activity limitations, bed days, sickness, chronic health problems, pain, and worry. Favorable health perceptions were related to psychological well-being and, to a marked degree, favorable health ratings obtained using a different method. Current Health scores were substantially related to both physical and mental health variables. These results constitute substantial support for the validity of HPQ scales as general

[2] All of these explanations of score differences should be kept in mind when results based on this scale are interpreted. It is recommended that the summary of factor structures across field tests (see Table 17) be used as an adjunct to other validity findings in interpreting HPQ scale scores.

measures of health. Studies of associations between general health ratings and other health status measures published before construction of the HPQ also supported this interpretation.

Although this evidence increased understanding of the meaning of general health ratings, little is known about their clinical significance. HIS data (Stewart, Ware, Brook, 1977; Vol. VI) suggest, particularly when nonlinear relationships with health status variables were examined, that (1) many very low general health rating scores (those in the lowest decile of the distribution) reflect chronic functional limitations due to poor health, chronic diseases, and psychiatric impairment; and (2) very high general health rating scores reflect states of positive well-being (e.g., vigor and happiness). Considerable effort beyond these studies will be necessary to determine what, if any, clinical significance can be assigned to general health rating scores that fall in the middle of the distribution.

Health and Illness Behavior. Studies of the validity of HPQ scales and most other general health ratings in relation to health and illness behavior used cross-sectional research designs. Ratings of health (at a given point) were correlated with information obtained at the same time about *previous* health and illness behavior. These data were used to test validity hypotheses regarding the influence of health perceptions on subsequent behavior. Despite the limitations of cross-sectional data, hypotheses regarding associations between general health ratings and health and illness behavior were confirmed, supporting the validity of such ratings as measures of health.

Correlations between illness behavior and most HPQ scales were significant in the hypothesized direction in all field tests and often substantial (e.g., Current Health in relation to doctor visits). Findings regarding the negative relationship between general health perceptions and use of health care services were consistent with those reported by others (Friedsam and Martin, 1963; Fabrega and Roberts, 1972; Tissue, 1972). Only Current Health and Health-Worry/Concern were significantly related to whether respondents reported compliance with medical regimens. Consistent with study hypotheses, previous findings for an earlier version of the HPQ (Ware, Wright, and Snyder, 1974), and results reported by Tissue (1972), health perceptions were not related to health behavior (e.g., checkups). Current Health and Prior Health scales were more strongly and more consistently (across field tests) correlated with illness behavior than were other HPQ health scales. This trend may have been an artifact of the cross-sectional study design. A test of the predictive validity of HPQ scales using a prospective study design and carefully controlling for differences in scale reliability would make it possible to determine which scales best predict use of health care services.

With one exception, correlations among the HPQ scales were not taken into account when their validity was evaluated in relation to behavior, so little is known about whether all of the scales make a unique contribution to prediction. This important issue was addressed by preliminary multiple linear regression analyses. Findings reported by Ware and Karmos (1976) and additional analyses of their data indicated that several of the HPQ scales make a unique contribution to the explanation of health and illness behavior. When three health and illness behaviors were regressed on eight[3] HPQ scales, the scales explained approximately 15 to 20 percent

---

[3] Ware and Karmos (1976) included the six HPQ health scales and two scales to measure patient role propensity in their model.

of variance in doctor visits, 2 to 6 percent of variance in checkups, and 3 percent of variance in dental visits. For behavior studied, only three or four scales (not the same ones) were necessary to explain all of the variance that could be explained. This suggests that more than one HPQ scale may contribute to the prediction of health and illness behavior.

V. HEALTH INSURANCE STUDY MEASURES OF GENERAL HEALTH PERCEPTIONS

HEALTH INSURANCE STUDY HPQ

The 32 HPQ items that met factor analytic and multitrait scaling criteria in studies reported by Ware and Karmos (1976) were selected for use in HIS health questionnaires. These items were not included in the enrollment Medical History Questionnaire (MHQ) in Dayton, Ohio (the first HIS site) but they were included in the other three sites. They are also administered annually in Health Questionnaires and will be included in exit MHQs in all sites.

A copy of the HPQ used in the HIS is presented in App. B. Items and response categories are the same as in previous HPQs studied; questionnaire instructions and format have been improved.[1] Instructions have been shortened and simplified, response categories have been printed to the right of each item with columnar headings on both pages, and the questionnaire has been set in type that makes it easier to read.

Scoring rules for HIS-HPQ items are presented in App. C, and scoring instructions for scales in App. D. These scoring algorithms are based on previous research and will be tested before implementation in the HIS (see discussion below).

PLAN OF ANALYSIS

Before use of HPQ scales to test hypotheses about the effects of health care financing and use of health care services on health status, analyses of HIS data will be performed to increase understanding of HPQ scores in terms of reliability, stability over one- to five-year intervals, validity as measures of health status and as predictors of health and illness behavior, and precision (power) for purposes of testing HIS hypotheses. Tradeoffs involved in reducing the number of scale scores necessary to define general health perceptions, including construction of a health perceptions index, will also be explored.

Priorities for HIS studies of the HPQ general health perceptions measures are based on the results and shortcomings of previous studies. They fall into three categories: (1) research design, (2) comprehensiveness of criterion variables, and (3) analytic methods. Two features of the HIS study design will greatly increase the strength of conclusions about the validity of general health ratings. First, longitudinal data from the HIS will permit prospective studies of the validity of general health ratings as predictors of consumption of health care services and of other health and illness behavior (e.g., change in providers, use of medication). Such studies are needed because it is impossible to establish these relationships with cross-sectional data.

Repeated administrations of the HPQ in the HIS will also allow validation of HPQ scales that differ with respect to time (perceptions regarding past, current,

and future health). For example, it will be possible to determine whether people who have been ill and who are now well score low on Prior Health and high on Current Health, and whether people scoring low on Prior Health, Current Health, and Health Outlook have chronic illnesses. Personal profiles of HPQ scores may be more clinically relevant than scores interpreted separately or in an additive model. Analyses of such profiles may permit inferences about the direction and extent of changes in health status over time using cross-sectional data. For example, it will be possible to test the hypothesis that people scoring high on Prior Health and low on Current Health (at any point) have experienced deterioration in health status (over time).

The second HIS design feature pertains to the sources of information about individual health status. Although a great deal remains to be learned about the meaning of general health ratings from studies in which all data come from self-report, the validity of general health ratings must be studied in relation to data obtained from sources other than the respondent (e.g., physicians, laboratory tests). Data from multiple sources would make it possible to determine the nature and number of data elements and sources necessary for a comprehensive definition of health and to determine the validity of general health ratings in relation to them. For example, HIS data will permit analysis of the validity of general health ratings in relation to physician diagnoses, the results of comprehensive screening exams performed by trained technicians, and the extent of disability reported in biweekly diaries.

The comprehensiveness of health status data gathered from HIS participants will make it possible to determine the extent to which ratings measure both objective (e.g., work loss days) and subjective (e.g., pain) health status constructs, and measure specific components of health (e.g., physical, mental, and social well-being). Preliminary findings based on general health rating scales other than the HPQ and data gathered in the Dayton HIS site are reported in Vol. VI. Studies of differences in data-gathering and scaling methods will determine the effects of method on conclusions about health status. Studies of the effects of method on patient satisfaction ratings (Ware, Snyder, and Wright, 1976) suggest that considerable method effects operate in survey instruments. The ability to understand validity is greatly limited by inability to control these effects on validity coefficients. Finally, with respect to the comprehensiveness of HIS data files, it will be possible to study nonhealth-related variables that may correlate with health status scores. For example, associations between measures of patient role propensity and health status can be studied. Evidence that scores from instruments based on different measurement methods are correlated is not sufficient; it is also essential to demonstrate that general health ratings and other health status measures do not correlate with variables other than health status measured using the same methods.

In addition, several strategies will be used to explore the construction and validation of a general health perceptions index. One approach would be to construct a single health perceptions index based on the first unrotated factor underlying correlations among HPQ items or scales. A second approach would be to measure only the perceptual dimension(s) with a high probability of usefulness to the goal of a given analysis. The validity studies reported here may help in selecting from among the HPQ scales those likely to be most useful in testing hypotheses about the effects of health care financing on health status.

Statistical considerations will also be important. The HIS will use multivariate statistical methods to take correlations among health constructs into account. It may not be sufficient to measure only the health construct relevant to a particular hypothesis. Health constructs are substantially interrelated both in theory and in measurement practice. Therefore, it is necessary to measure constructs that may not be directly relevant to interpret study findings with any degree of confidence.

Interactions among health constructs and between health constructs and other health care variables and respondent characteristics will also be considered. Preliminary findings reported by Ware, Snyder, and Wright (1976) indicate, for example, that the relationship between health status and use of health care services is not the same at all levels of patient satisfaction with medical care. Further, Johnston and Ware (1976) have shown that the relationship between general health ratings and physical and mental health constructs may differ for upper and lower income groups. Studies of population differences in the meaning of general health ratings along these lines will be conducted to gain a better understanding of how general health ratings should be used and of their validity across population groups.

HIS studies will also address problems of response set, which have been ignored in many health status surveys. Response biases—tendencies to endorse or negate items regardless of content or to respond in a socially desirable manner—may be a problem in standardized survey measures of health and health care perceptions (Ware and Karmos, 1976; Ware, 1978). These response sets, which are related to the method effects discussed earlier, may bias estimates of reliability, relationships among health constructs, validity coefficients, and population differences in health status.

Scales to measure general health perceptions yield considerable information about health status. The level of measurement and nature of scale distributions on these scales make them suitable for use with the more powerful statistical methods (e.g., multiple linear regression) to test hypotheses regarding effects of differences in health care financing on health status. If results of these tests are to be interpretable, however, the validity of HPQ scale scores (their meaning) must be well understood. For these reasons, additional studies of validity using HIS data will precede the testing of mainline HIS hypotheses.

Appendix A

SUPPORTING TABLES

Table A.1

INTERNAL-CONSISTENCY RELIABILITY COEFFICIENTS
FOR HPQ FORM II HEALTH SCALES,
FOUR FIELD TESTS

Scales	No. of Items	Field Tests[a]			
		SAC	ESL	FPC	LAC
Current Health	9	.91	.89	.92	.92
Prior Health	3	.70	.70	.79	.75
Resistance-Susceptibility	4	.70	.58	.80	.72
Health Outlook	4	.74	.64	.79	.76
Health Worry/Concern	4	.53	.45	.61	.62
Sickness Orientation	2	.58	.46	.54	.61

[a]SAC = Sangamon County, ESL = East St. Louis,
FPC = Family Practice Center, LAC = Los Angeles
County.

Table A.2

CORRELATIONS BETWEEN HPQ HEALTH SCALES AND FOUR ROTATED FACTORS,
SANGAMON COUNTY FIELD TEST (N=432)
(Highest loading in each row underlined)

Scales	Factors				Proportion of Variance Explained
	Prior Health	Current Health/ Health Outlook	Health Worry/ Concern	Sickness Orientation	
Prior Health	.77	.24	.14	-.09	.68
Current Health	.42	.78	.19	-.06	.82
Health Outlook	.15	.75	.21	-.17	.66
Health Worry/Concern	-.14	-.11	-.66	.08	.47
Sickness Orientation	-.07	-.12	-.12	.74	.58
Resistance-Susceptibility	.04	.46	.59	-.18	.60

Table A.3

CORRELATIONS BETWEEN HPQ HEALTH SCALES AND FOUR ROTATED FACTORS,
EAST ST. LOUIS FIELD TEST (N=323)
(Highest loading in each row underlined)

	Factors				
Scales	Prior Health	Current Health/ Health Outlook	Health Worry/ Concern	Sickness Orientation	Proportion of Variance Explained
Prior Health	.80	.10	.07	-.12	.67
Current Health	.60	.64	.19	-.06	.82
Health Outlook	.11	.73	.12	-.20	.60
Health Worry/Concern	-.10	-.11	-.64	.05	.44
Sickness Orientation	-.12	-.14	-.06	.64	.45
Resistance-Susceptibility	.51	.29	.34	-.26	.52

Table A.4

CORRELATIONS BETWEEN HPQ HEALTH SCALES AND FOUR ROTATED FACTORS,
FAMILY PRACTICE FIELD TEST (N=520)
(Highest loading in each row underlined)

	Factors				
Scales	Prior Health	Current Health/ Health Outlook	Health Worry/ Concern	Sickness Orientation	Proportion of Variance Explained
Prior Health	.83	.24	.17	-.10	.78
Current Health	.31	.82	.28	-.04	.84
Health Outlook	.11	.80	.13	-.18	.70
Health Worry/Concern	-.14	-.19	-.72	.09	.58
Sickness Orientation	-.05	-.08	-.06	.69	.49
Resistance-Susceptibility	.23	.45	.40	-.45	.62

Table A.5

CORRELATIONS BETWEEN HPQ HEALTH SCALES AND FOUR ROTATED FACTORS,
LOS ANGELES COUNTY FIELD TEST (N=515)
(Highest loading in each row underlined)

	Factors				
Scales	Prior Health	Current Health/ Health Outlook	Health Worry/ Concern	Sickness Orientation	Proportion of Variance Explained
Prior Health	.82	.24	-.10	-.10	.74
Current Health	.32	.84	-.23	.00	.86
Health Outlook	.11	.77	-.09	-.29	.70
Health Worry/Concern	-.06	-.14	.74	.11	.58
Sickness Orientation	-.08	-.16	.14	.74	.59
Resistance-Susceptibility	.23	.48	-.42	-.26	.52

Table A.6

CORRELATIONS BETWEEN PRIOR HEALTH AND HEALTH-RELATED
VARIABLES AND AGE, FOUR FIELD TESTS

	Field Tests			
Validity Variables	East St. Louis	Sangamon County	Family Practice	Los Angeles County
Health Variables				
General health item	$--^a$	--	$.50^d$	--
Role activity limitations[b]	$-.38^d$	$-.35^d$	$-.42^d$	$-.28^d$
Any bed days	--	--	--	$-.21^d$
Total bed days	--	--	--	$-.30^{c,d}$
Sickness	--	--	--	$-.18^d$
Chronic problems	--	--	--	$-.32^d$
Pain	--	--	$-.39^d$	--
Feeling poorly	--	--	--	$-.19^d$
Health worry	--	--	$-.44^d$	--
Psychological well-being	--	--	--	$.19^d$
Reason for last doctor visit	--	$.11^d$	--	--
Health and Illness Behavior				
Number of doctor visits	$-.32^d$	$-.26^d$	$-.35^d$	--
Dental visits	-.02	-.10	$.11^d$	--
Check-up	-.08	-.03	.07	--
Doctor visit when needed	--	--	--	$-.27^d$
Recency of care	--	$.16^d$	--	$.14^d$
Hospitalization	$-.28^d$	$-.28^d$	$-.26^d$	--
Compliance	--	--	--	-.07
Patient Role Propensity				
Rejection of Sick Role	.01	.01	-.03	-.06
Attitude Toward Going to the Doctor	-.01	.01	$.13^d$	.00
Sociodemographic Variable				
Age	$-.16^d$	$-.25^d$	$-.19^d$	$-.20^d$

[a]-- indicates not tested.

[b]Role Activity Limitations was scored as a dichotomous variable in the Los Angeles County field test and as a continuous variable in all other field tests.

[c]$N = 225$.

[d]$p < 0.01$, one-tailed test.

Table A.7

Correlations Between Current Health and Health-Related Variables and Age, Four Field Tests

Validity Variables	Field Tests			
	East St. Louis	Sangamon County	Family Practice	Los Angeles County
Health Variables				
General health item	--[a]	--	.80[c]	--
Role activity limitations[b]	-.47[c]	-.46[c]	-.45[c]	-.37[c]
Any bed days	--	--	--	-.37[c]
Total bed days	--	--	--	-.38[c]
Sickness	--	--	--	-.24[c]
Chronic problems	--	--	--	-.43[c]
Pain	--	--	-.55[c]	--
Feeling poorly	--	--	--	-.28[c]
Health worry	--	--	-.58[c]	--
Psychological well-being	--	--	--	.39[c]
Reason for last doctor visit	--	.21[c]	--	--
Health and Illness Behavior				
Number of doctor visits	-.42[c]	-.32[c]	-.44[c]	--
Dental visits	.01	-.15[c]	.03	--
Check-up	-.05	-.13[c]	.08	--
Doctor visit when needed	--	--	--	-.47[c]
Recency of care	--	.10	--	.22[c]
Hospitalization	-.25[c]	-.30[c]	-.36[c]	--
Compliance	--	--	--	-.14[c]
Patient Role Propensity				
Rejection of Sick Role	-.03	.04	-.02	-.01
Attitude Toward Going to the Doctor	.03	.13[c]	.22[c]	.05
Sociodemographic Variable				
Age	-.26[c]	-.32[c]	-.20[c]	-.26[c]

[a] -- indicates not tested.

[b] Role Activity Limitations was scored as a dichotomous variable in the Los Angeles County field test and as a continuous variable in all other field tests.

[c] $p < 0.01$, one-tailed test.

Table A.8

Correlations Between Health Outlook and Health-Related Variables and Age, Four Field Tests

Validity Variables	Field Tests			
	East St. Louis	Sangamon County	Family Practice	Los Angeles County
Health Variables				
General health item	--[a]	--	.58[c]	--
Role activity limitations[b]	-.24[c]	-.23[c]	-.24[c]	-.19[c]
Any bed days	--	--	--	-.23[c]
Total bed days	--	--	--	-.16[c]
Sickness	--	--	--	-.13[c]
Chronic problems	--	--	--	-.34[c]
Pain	--	--	-.35[c]	--
Feeling poorly	--	--	--	-.13[c]
Health worry	--	--	-.33[c]	--
Psychological well-being	--	--	--	.28[c]
Reason for last doctor visit	--	.09	--	--
Health and Illness Behavior				
Number of doctor visits	-.16[c]	-.14[c]	-.20[c]	--
Dental visits	.06	-.09	.06	--
Check-up	-.01	-.04	.11[c]	--
Doctor visit when needed	--	--	--	-.26[c]
Recency of care	--	.05	--	.08
Hospitalization	-.07	-.11[c]	-.21[c]	--
Compliance	--	--	--	-.06
Patient Role Propensity				
Rejection of Sick Role	.10	.03	.05	.06
Attitude Toward Going to the Doctor	-.04	.08	.15[c]	.06
Sociodemographic Variable				
Age	-.17[c]	-.26[c]	-.20[c]	-.28[c]

[a]-- indicates not tested.

[b]Role Activity Limitations was scored as a dichotomous variable in the Los Angeles field test and as a continuous variable in all other field tests.

[c]$p < 0.01$, one-tailed test.

Table A.9

CORRELATIONS BETWEEN RESISTANCE-SUSCEPTIBILITY AND
HEALTH-RELATED VARIABLES AND AGE, FOUR FIELD TESTS

	Field Tests			
Validity Variables	East St. Louis	Sangamon County	Family Practice	Los Angeles County
Health Variables				
General health item	--[a]	--	.50	--
Role activity limitations[b]	-.37[c]	-.29[c]	-.39[c]	-.28[c]
Any bed days	--	--	--	-.24[c]
Total bed days	--	--	--	.00
Sickness	--	--	--	-.21[c]
Chronic problems	--	--	--	-.26[c]
Pain	--	--	-.41[c]	--
Feeling poorly	--	--	--	-.21[c]
Health worry	--	--	-.42[c]	--
Psychological well-being	--	--	--	.18[c]
Reason for last doctor visit	--	.17[c]	--	--
Health and Illness Behavior				
Number of doctor visits	-.32[c]	-.17[c]	-.34[c]	--
Dental visits	-.02	-.02	.10	--
Check-up	-.01	-.02	-.01	--
Doctor visit when needed	--	--	--	-.21[c]
Recency of care	--	.12[c]	--	.19[c]
Hospitalization	-.26[c]	-.12[c]	-.27[c]	--
Compliance	--	--	--	-.06
Patient Role Propensity				
Rejection of Sick Role	.10	.19[c]	.12[c]	.20[c]
Attitude Toward Going to the Doctor	-.05	.06	.14[c]	.00
Sociodemographic Variable				
Age	.09	.06	.09	.02

[a] -- indicates not tested.

[b] Role Activity Limitations was scored as a dichotomous variable in the Los Angeles County Field test and as a continuous variable in all other field tests.

[c] $p < 0.01$, one-tailed test.

Table A.10

CORRELATIONS BETWEEN HEALTH WORRY/CONCERN AND HEALTH-RELATED
VARIABLES AND AGE, FOUR FIELD TESTS

Validity Variables	Field Tests			
	East St. Louis	Sangamon County	Family Practice	Los Angeles County
Health Variables				
General health item	--[a]	--	-.38[c]	--
Role activity limitations[b]	.24[c]	.16[c]	.17[c]	.20[c]
Any bed days	--	--	--	.10[c]
Total bed days	--	--	--	.07
Sickness	--	--	--	.12[c]
Chronic problems	--	--	--	.10[c]
Pain	--	--	.24[c]	--
Feeling poorly	--	--	--	.14[c]
Health worry	--	--	.42[c]	--
Psychological well-being	--	--	--	-.12[c]
Reason for last doctor visit	--	-.05	--	--
Health and Illness Behavior				
Number of doctor visits	.20[c]	.15[c]	.29[c]	--
Dental visits	-.08	.03	-.01	--
Check-up	-.04	-.06	-.01	--
Doctor visit when needed	--	--	--	.17[c]
Recency of care	--	-.14[c]	--	-.15[c]
Hospitalization	-.07	-.12[c]	.08	--
Compliance	--	--	--	.16[c]
Patient Role Propensity				
Rejection of Sick Role	-.17[c]	-.29[c]	-.12[c]	-.24[c]
Attitude Toward Going to the Doctor	.16[c]	.04	.02	.00
Sociodemographic Variable				
Age	-.10	.07	-.09	-.11[c]

[a] -- indicates not tested.

[b] Role Activity Limitations was scored as a dichotomous variable in the Los Angeles County field test and as a continuous variable in all other field tests.

[c] $p < 0.01$, one-tailed test.

Table A.11

CORRELATIONS BETWEEN SICKNESS ORIENTATION AND HEALTH-RELATED
VARIABLES AND AGE, FOUR FIELD TESTS

Validity Variables	Field Tests			
	East St. Louis	Sangamon County	Family Practice	Los Angeles County
Health Variables				
General health item	--[a]	--	-.16[c]	--
Role activity limitations[b]	.11	.15[c]	.16[c]	.13[c]
Any bed days	--	--	--	.14[c]
Total bed days	--	--	--	.01
Sickness	--	--	--	.12[c]
Chronic problems	--	--	--	.14[c]
Pain	--	--	.18[c]	--
Feeling poorly	--	--	--	.09
Health worry	--	--	.20[c]	--
Psychological well-being	--	--	--	-.11[c]
Reason for last doctor visit	--	.01	--	--
Health and Illness Behavior				
Number of doctor visits	.11	.08	.10[c]	--
Dental visits	.00	.01	-.04	--
Check-up	.02	.03	-.05	--
Doctor visit when needed	--	--	--	.09
Recency of care	--	-.01	--	.01
Hospitalization	.10	.01	.02	--
Compliance	--	--	--	-.10
Patient Role Propensity				
Rejection of Sick Role	.00	-.09	-.11[c]	-.07
Attitude Toward Going to the Doctor	.06	-.01	.00	-.05
Sociodemographic Variable				
Age	.04	-.03	-.13[c]	-.08

[a] -- indicates not tested.

[b] Role Activity Limitations was scored as a dichotomous variable in the Los Angeles County field test and as a continuous variable in all other field tests.

[c] $p < 0.01$, one-tailed test.

Appendix B
HEALTH PERCEPTIONS QUESTIONNAIRE

128. PLEASE READ EACH OF THE FOLLOWING STATEMENTS, AND THEN CIRCLE ONE OF THE NUMBERS <u>ON EACH LINE</u> TO INDICATE WHETHER THE STATEMENT IS TRUE OR FALSE <u>FOR YOU</u>.

<u>THERE ARE NO RIGHT OR WRONG ANSWERS</u>.

If a statement is definitely true for you, circle 5.
If it is mostly true for you, circle 4.
If you don't know whether it is true or false, circle 3.
If it is mostly false for you, circle 2.
If it is definitely false for you, circle 1.

SOME OF THE STATEMENTS MAY LOOK OR SEEM LIKE OTHERS. BUT EACH STATEMENT IS DIFFERENT, AND SHOULD BE RATED BY ITSELF.

	Definitely true	Mostly true	Don't know	Mostly false	Definitely false	
A. According to the doctors I've seen, my health is now excellent	5	4	3	2	1	26/
B. I try to avoid letting illness interfere with my life	5	4	3	2	1	27/
C. I seem to get sick a little easier than other people	5	4	3	2	1	28/
D. I feel better now than I ever have before	5	4	3	2	1	29/
E. I will probably be sick a lot in the future	5	4	3	2	1	30/
F. I never worry about my health	5	4	3	2	1	31/
G. Most people get sick a little easier than I do	5	4	3	2	1	32/
H. I don't like to go to the doctor	5	4	3	2	1	33/
I. I am somewhat ill	5	4	3	2	1	34/
J. In the future, I expect to have better health than other people I know	5	4	3	2	1	35/
K. I was so sick once I thought I might die	5	4	3	2	1	36/
L. I'm not as healthy now as I used to be	5	4	3	2	1	37/
M. I worry about my health more than other people worry about their health	5	4	3	2	1	38/

		Definitely true	Mostly true	Don't know	Mostly false	Definitely false	
N.	When I'm sick, I try to just keep going as usual	5	4	3	2	1	39/
O.	My body seems to resist illness very well	5	4	3	2	1	40/
P.	Getting sick once in a while is a part of my life	5	4	3	2	1	41/
Q.	I'm as healthy as anybody I know	5	4	3	2	1	42/
R.	I think my health will be worse in the future than it is now	5	4	3	2	1	43/
S.	I've never had an illness that lasted a long period of time	5	4	3	2	1	44/
T.	Others seem more concerned about their health than I am about mine	5	4	3	2	1	45/
U.	When I'm sick, I try to keep it to myself	5	4	3	2	1	46/
V.	My health is excellent	5	4	3	2	1	47/
W.	I expect to have a very healthy life	5	4	3	2	1	48/
X.	My health is a concern in my life	5	4	3	2	1	49/
Y.	I accept that sometimes I'm just going to be sick	5	4	3	2	1	50/
Z.	I have been feeling bad lately	5	4	3	2	1	51/
AA.	It doesn't bother me to go to a doctor	5	4	3	2	1	52/
BB.	I have never been seriously ill	5	4	3	2	1	53/
CC.	When there is something going around, I usually catch it	5	4	3	2	1	54/
DD.	Doctors say that I am now in poor health	5	4	3	2	1	55/
EE.	When I think I am getting sick, I fight it	5	4	3	2	1	56/
FF.	I feel about as good now as I ever have	5	4	3	2	1	57/

Appendix C

SCORING RULES FOR HIS-HPQ ITEMS

Scoring	Items
5=Definitely true 4=Mostly true 3=Don't know 2=Mostly false 1=Definitely false	A, B, D, G, J, M, N, O, P, Q, S, T, U, V, W, X, Y, AA, BB, EE, FF
1=Definitely true 2=Mostly true 3=Don't know 4=Mostly false 5=Definitely false	C, E, F, H, I, K, L, R, Z, CC, DD

Appendix D

SCORING RULES FOR HIS-HPQ SCALES

Scale	Sum Scores for These Items
Prior Health	K + S + BB
Current Health	A + D + I + L + Q + V + Z + DD + FF
Health Outlook	E + J + R + W
Health Worry/Concern	F + M + T + X
Resistance-Susceptibility	C + G + O + CC
Sickness Orientation	P + Y
Rejection of Sick Role	B + N + U + EE
Attitude Toward Going to the Doctor	H + AA

BIBLIOGRAPHY

Bem, D. J., "Self-Perception Theory," in L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Academic Press, New York, 1972.

Bergner, M., R. A. Bobbitt, W. E. Pollard, et al., "The Sickness Impact Profile: Validation of a Health Status Measure," *Medical Care* 14:57–67, 1976.

Berkman, P., "Life Stress and Psychological Well-Being: A Replication of Langner's Analysis in the Midtown Manhattan Study," *Journal of Health and Social Behavior* 12:35–45, 1971.

Bradburn, N. M., *The Structure of Psychological Well-Being*, Aldine Publishing Co., Chicago, 1969.

Campbell, D. T. and D. W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin* 56:81–105, 1959.

Cavan, R. S., G. W. Burgess, R. J. Havighurst, et al., *Personal Adjustment to Old Age*, Science Research Associates, Chicago, 1949.

Cole, S. and R. Lejeune, "Illness and the Legitimation of Failure," *American Sociological Review* 37:347–356, 1972.

Comrey, A. L., "Tandem Criteria for Analytic Rotation in Factor Analysis," *Psychometrika* 16:297–334, 1967.

Crawford, M. P. and D. Hooper, "Menopause, Aging and Family," *Social Science and Medicine* 7:469–482, 1973.

Cronbach, L. J., "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika* 16:297–334, 1951.

Cutler, S., "Voluntary Association Participation and Life Satisfaction: A Cautionary Research Note," *Journal of Gerontology* 28:96–100, 1973.

Desroches, H. F., B. D. Kaiman, and H. T. Ballard, "Factors Influencing Reporting of Physical Symptoms by the Aged Patient," *Geriatrics* 22:169–175, 1967.

DiCicco, L. and D. Apple, "Health Needs and Opinions of Older Adults," *Public Health Reports* 73:479–487, 1958.

Dupuy, H. J., "The Psychological Section of the Current Health and Nutrition Examination Survey," *Proceedings of the Public Health Conference on Records and Statistics Meeting Jointly with the National Conference on Health Statistics*, National Center for Health Statistics, Washington, D.C., June 1972.

Edwards, J. and D. Klemmack, "Correlates of Life Satisfaction: A Reexamination," *Journal of Gerontology* 28:497–502, 1973.

Fabrega, H. and R. E. Roberts, "Social-Psychological Correlates of Physician Use by Economically Disadvantaged Negro Urban Residents," *Medical Care* 10:215–223, 1972.

Fine, M., "Interrelationships Among Mobility, Health, and Attitudinal Variables in an Urban Elderly Population," *Human Relations* 28:451–474, 1975.

Friedsam, H. H. and H. W. Martin, "A Comparison of Self and Physicians' Health Ratings in an Older Population," *Journal of Health and Human Behavior* 4:179, 1963.

Gaitz, C. M. and J. Scott, "Age and the Measurement of Mental Health," *Journal of Health and Social Behavior* 12:55–67, 1972.

Garrity, T. F., "Vocational Adjustment After First Myocardial Infarction: Comparative Assessment of Several Variables Suggested in the Literature," *Social Science and Medicine* 7:705–717, 1973.

Guttman, L. A., "A Basis for Scaling Qualitative Data," *American Sociological Review* 9:139, 1944.

Helmstadter, G. C., *Principles of Psychological Measurement*, Appleton-Century-Crofts, New York, 1964.

Henley, B. and M. S. Davis, "Satisfaction and Dissatisfaction: A Study of the Chronically-Ill Aged Patient," *Journal of Gerontology* 18:65–75, 1967.

Heyman, D. K. and F. C. Jeffers, "Effect of Time Lapse on Consistency of Self-Health and Medical Evaluations of Elderly Persons," *Journal of Gerontology* 18:160–164, 1963.

Hochstim, J., D. Athanasopoulos, and J. Larkins, "Poverty Area Under the Microscope," *American Journal of Public Health* 58:1815–1827, 1968.

Johnston, S. A. and J. E. Ware, Jr., "Income Group Differences in Relationships Among Survey Measures of Physical and Mental Health," *Health Services Research* 11:416–429, 1976.

Kessel, N. and M. Shepard, "The Health and Attitudes of People who Seldom Consult a Doctor," *Medical Care* 3:6, 1965.

Leveson, I., "Access to Medical Care: The Queensbridge Experiment," *Inquiry* 9:61–68, 1972.

Likert, R., "A Technique for the Measurement of Attitudes," *Archives of Psychology* 140:1–55, 1932.

Ludwig, E. and G. Gibson, "Self Perception of Sickness and the Seeking of Medical Care," *Journal of Health and Social Behavior* 10:125, 1969.

MacPherson, A. S., "The Measurement of Mental Health in a General Population," thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science in Medicine, McMaster University, Hamilton, Ontario, 1972.

Maddox, G. L., "Activity and Morale: A Longitudinal Study of Selected Elderly Subjects," *Social Forces* 42:195–204, 1963.

———, "Self-Assessment of Health Status: A Longitudinal Study of Selected Elderly Subjects," *Journal of Chronic Diseases* 17:449–460, 1964.

——— and C. Eistorfer, "Some Correlates of Activity and Morale Among the Elderly" *Social Forces* 40:254, 1962.

Madow, W. G., "Interview Data on Chronic Conditions Compared with Information Derived from Medical Records," *Vital and Health Statistics* Series 2, No. 23, National Center for Health Statistics, Rockville, Md., 1967.

Maiman, L. A. and M. H. Becker, "The Health Belief Model: Origins and Correlates in Psychological Theory," *Health Education Monographs* 2:336–353, 1974.

Maiman, L. A., M. H. Becker, J. P. Kirscht, et al., "Scales for Measuring Health Belief Model Dimensions: A Test of Predictive Value, Internal Consistency, and Relationships Among Beliefs," *Health Education Monographs* 5:215–230, 1977.

National Center for Health Statistics, *Health: United States, 1975*, DHEW Publication No. HRA 76-1232, U.S. Department of Health, Education, and Welfare, Rockville, Md., 1976, p. 243.

Nowlin, J. B., "Depression and Health," in E. Palmore (ed.), *Normal Aging, II*, Duke University Press, Durham, North Carolina, 1974; pp. 168-172.
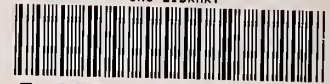
Osborn, R. W., "Social Rank and Self-Health Evaluation of Older Urban Males," *Social Science and Medicine* 7:209–218, 1973.

Palmore, E. and C. Luikart, "Health and Social Factors Related to Life Satisfaction," *Journal of Health and Social Behavior* 13:68, 1972.

Pratt, L., "The Relationship of Socioeconomic Status to Health," *American Journal of Public Health* 61:281, 1971.

——, "Conjugal Organization and Health," *Journal of Marriage and the Family* 34:85–95, 1972.

Reeder, L. G., L. G. Ramacher, S. Gorelnik, *Handbook of Scales and Indices of Health Behavior*, Goodyear Publishing Co., Pacific Palisades, Ca., 1976.

Renne, K. S., "Correlates of Dissatisfaction in Marriage," *Journal of Marriage and the Family* 22:54, 1970.

——, "Health and Marital Experience in an Urban Population," *Journal of Marriage and the Family* 23:338, 1971.

——, "Measurement of Social Health in a General Population Survey," *Social Science Research* 3:25, 1974.

Reynolds, J., W. A. Rushing, and D. L. Miles, "Validation of a Function Status Index," *Journal of Health and Social Behavior* 15:271, 1974.

Rushing, W., Personal Communication, 1977.

Schnore, L. F. and J. D. Cowhig, "Some Correlates of Reported Health in Metropolitan Centers," *Social Problems* 7:218, 1959-1960.

Shanas, E., P. Townsend, D. Weddarburn, et al., *Old People in Three Industrial Societies*, Atherton Press, New York, 1968, pp. 49–70.

Spreitzer, E. and E. E. Snyder, "Correlates of Life Satisfaction Among the Elderly," *Journal of Gerontology* 29:454, 1974.

Stewart, A. L., J. E. Ware, Jr., and R. H. Brook, "The Meaning of Health: Understanding Functional Limitations," *Medical Care* 15:939–953, 1977.

Suchman, E. A., B. S. Phillips, and G. E. Streib, "An Analysis of the Validity of Health Questionnaires," *Social Forces* 26:223, 1958.

Thompson, W. E. and G. F. Streib, "Situational Determinants: Health and Economic Deprivation in Retirement," *Journal of Social Issues* 14:18–45, 1958.

Tissue, T., "Another Look at Self-Rated Health Among the Elderly," *Journal of Gerontology* 27:91, 1972.

Tornstam, L., "Health and Self-Perception: A Systems Theoretical Approach," *Gerontologist* 15:264–270, 1975.

Ware, J. E., Jr., "Scales to Measure General Health Perceptions," *Health Services Research* 11:396–415, 1976.

——, "Effects of Acquiescent Response Set on Patient Satisfaction Ratings," *Medical Care* 16:327–336, 1978.

Ware, J. E., Jr. and A. H. Karmos, *Development and Validation of Scales to Measure Perceived Health and Patient Role Propensity: Volume II of a Final Report,* Southern Illinois University School of Medicine, Carbondale, Ill., 1976.

Ware, J. E., Jr., M. K. Snyder, and W. R. Wright, *Development and Validation of Scales to Measure Patient Satisfaction with Health Care Services: Volume I of a Final Report.* "Part B: Results Regarding Scales Constructed from the Patient Satisfaction Questionnaire and Measures of Other Health Perceptions," Southern Illinois University School of Medicine, Carbondale, Illinois, 1976.

Ware, J. E., Jr., W. R. Wright, and M. K. Snyder, *Measures of Perception Regarding Health Status: Preliminary Findings*, Publication No. PB-242-726/AS, National Technical Information Service, Springfield, Virginia, 1974.

Warren, H. H., "Self-Perception of Independence Among Urban Elderly," *American Journal of Occupational Therapy* 28:329–336, 1974.